

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Transcriptional profiling of subcutaneous adipose tissue in Italian Large White pigs divergent for backfat thickness

This is the submitted version (pre peer-review, preprint) of the following publication:

Published Version:

Zambonelli, P., Gaffo, E., Zappaterra, M., Bortoluzzi, S., Davoli, R. (2016). Transcriptional profiling of subcutaneous adipose tissue in Italian Large White pigs divergent for backfat thickness. *ANIMAL GENETICS*, 47(3), 306-323 [10.1111/age.12413].

Availability:

This version is available at: <https://hdl.handle.net/11585/581992> since: 2017-03-15

Published:

DOI: <http://doi.org/10.1111/age.12413>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

Transcriptional profiling of subcutaneous adipose tissue in Italian Large White pigs divergent for backfat thickness

Journal:	<i>Animal Genetics</i>
Manuscript ID	AnGen-15-06-0185.R2
Manuscript Type:	Full Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Zambonelli, Paolo; Bologna University, Department of Agricultural and-Food Sciences (DISTAL) Gaffo, Enrico; University of Padova, Department of Molecular Medicine Zappaterra, Martina; Bologna University, Department of Agricultural and-Food Sciences (DISTAL) Bortoluzzi, Stefania; University of Padova, Department of Molecular Medicine Davoli, Roberta; Bologna University, Department of Agricultural and-Food Sciences (DISTAL)
Keywords:	backfat, fat deposition, gene expression, differential analysis, pigs

This is the preprint version of the accepted article, then published in *Animal genetics*, Volume 47, Issue 3, Pages 306-323 which has been published in final form at DOI <https://doi.org/10.1111/age.12413>

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 **Transcriptional profiling of subcutaneous adipose tissue in Italian**
2 **Large White pigs divergent for backfat thickness**

3
4 Zambonelli Paolo^{1*}, Gaffo Enrico^{2*}, Zappaterra Martina¹, Bortoluzzi Stefania^{2§}, Davoli
5 Roberta^{1§}

6
7 ¹Department of Agricultural and-Food Sciences (DISTAL), Bologna University, Via Fratelli
8 Rosselli 107, 42123 Reggio Emilia, Italy.

9 ²Department of Molecular Medicine, University of Padova, Via Gabelli 63, 35121 Padova,
10 Italy

11
12 [§] To whom correspondence should be addressed. Roberta Davoli Tel. +39 051 2088050, Fax
13 +39 051 2086172, roberta.davoli@unibo.it; Stefania Bortoluzzi Tel: +39 049 8276502, Fax:
14 +39 049 8276209, stefania.bortoluzzi@unipd.it

15
16 *These authors contributed equally to this work

18 Summary

19 Fat deposition is a widely studied trait in pigs for the implications with animal growth
20 efficiency, technological and nutritional characteristics of meat products, but the global
21 framework of the biological and molecular processes regulating fat deposition in pigs is still
22 incomplete. This paper describes the backfat tissue transcription profile in Italian Large
23 White pigs and reports genes differentially expressed between fat and lean animals
24 according to RNA-seq data. The backfat transcription profile was characterized by the
25 expression of 23,483 genes of which 54.1% were represented by known genes. Of 63,418
26 expressed transcripts, about 80% were non previously annotated isoforms. By comparing
27 the expression level of fat vs. lean pigs we detected 86 robust differentially expressed
28 transcripts, 72 more expressed (e.g. *ACP5*, *BCL2A1*, *CCR1*, *CD163*, *CD1A*, *EGR2*, *ENPP1*,
29 *GPNMB*, *INHBB*, *LYZ*, *MSR1*, *OLR1*, *PIK3AP1*, *PLIN2*, *SPP1*, *SLC11A1*, *STC1*) and 14 less
30 expressed (e.g. *ADSSL1*, *CDO1*, *DNAJB1*, *HSPA1A*, *HSPA1B*, *HSPA2*, *HSPB8*, *IGFBP5*, *OLFML3*)
31 in fat pigs. The main functional categories enriched in differentially expressed genes were
32 immune system process, response to stimulus, cell activation, and skeletal system
33 development, for the overexpressed, unfolded protein binding and stress response, for the
34 under-expressed genes, which include five heat shock proteins. Adipose tissue alterations
35 and impaired stress response are linked to inflammation and, in turn, to adipose tissue
36 secretory activity similarly to what is observed in human obesity. Our results open the
37 opportunity to identify biomarkers of carcass fat traits to improve pig production chain and
38 to identify genetic factors that regulate the observed differential expression.

40 **Keywords:** backfat, fat deposition, gene expression, differential analysis, pigs

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

41

42

43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

Backfat deposition and fat traits are among the most important characters studied in pigs, due to their strong relation with human nutrition of pig products and for the technological characteristics of high quality Protected Designation of Origin (PDO) dry-cured hams. The amount of fat laid on the external part of the pig body (subcutaneous fat or backfat) is of extreme importance for growth performances, as the lesser is the deposited fat, the better the growth performances. Regarding technological aspects related to the dry-cured high quality products and meat industry, an adequate layer of fat is required for the seasoning process of PDO products, like dry cured hams (Bosi & Russo, 2004; Čandek-Potokar & Škrlep, 2012).

During the last decade, pig transcriptomic data have been obtained initially by expressed sequence tag sequencing (Mikawa et al., 2004; Uenishi et al., 2004; Chen et al., 2006; Gorodkin et al., 2007; Uenishi et al., 2007) and microarrays (Hornshøj et al., 2007; Ferraz et al., 2008; Moon et al., 2009), which allowed the comparison of gene expression level in several pig tissues. More recently, the RNA-seq approach was used to compare the transcription profile of different pig fat tissues or different pig breeds (Chen et al., 2011; Li et al., 2012; Corominas et al., 2013; Jiang et al., 2013; Zhou et al., 2013; Sodi et al., 2014; Toedebush et al., 2014; Wang et al. 2014). The differentially expressed genes (DEG) reported in these studies are useful to investigate the metabolic pathways activated by or associated with an increased fat deposition in pig body. However, the large amount of data produced and the results reported in literature are often hardly comparable because of differences in the studied breeds; heterogeneous animals' ages; and fat deposition stages.

Moreover, these researches identified several new genes and transcripts not reported in swine or other species. To date, the global framework of the biological processes regulating backfat deposition in pigs is still incomplete, and literature is poor of studies carried out on a homogeneous sample of individuals of the same breed reared on the same environmental conditions.

The objective of this research was to investigate the transcription profile of Italian Large White (ILW) pig backfat tissue and to compare the transcriptome of animals reared in the same herd and farming conditions and showing high (FAT) and low (LEAN) backfat thickness. Moreover a first functional characterization of DEGs has been obtained to provide new insights on genes, pathways and processes influencing the divergent aptitude of subcutaneous adipose tissue deposition in ILW pigs.

Materials and methods

Samples collection and RNA extraction

We sampled twenty individuals from a purebred population of 949 ILW sib-tested pigs provided by the Italian National Association of Pig Breeders (Associazione Nazionale Allevatori Suini, ANAS, <http://www.anas.it>. Accessed 22 June 2015). All animals used in this study were kept according to Italian and European law for pig production and all procedures described were in compliance with national and European Union regulations for animal care and slaughtering. The animals were reared on the ANAS Sib-Test genetic station from about 30 kg live weight to at least 155 kg live weight. For the genetic evaluation of a boar, full sib triplets (two females and one castrated male) were farmed on the genetic station to be performance tested. The formula and amount of the ration was the same for all. It was

1
2
3 87 based mainly on cereals and soybean, given in excess calculated using the “*quasi ad libitum*”
4
5 88 rule (a ration sufficiently abundant that 60% of pigs were able to ingest the full supplied
6
7
8 89 food). At the end of tests, animals were transported to a commercial abattoir located about
9
10 90 25 km far from the test station according to the Council Rule (EC) No 1/2005 on the
11
12 91 protection of animals during transport and related operations and amending Directives
13
14 92 64/432/EEC and 93/119/EC and Regulation (EC) No 1255/97. At slaughterhouse the pigs
15
16
17 93 were electrical stunned and bled in a lying position in agreement with the Council
18
19 94 Regulation (EC) No 1099/2009 on the protection of animals at the time of killing. All
20
21
22 95 slaughter procedures were monitored by the Veterinary team appointed by the Italian
23
24 96 Ministry of Health. Backfat samples were collected after slaughter, from 949 ILW pigs
25
26
27 97 slaughtered at an average hot carcass weight of 118.97 kg (± 0.29 SEM) and at an average
28
29 98 age of eight months during the years 2011 and 2012 in 27 different slaughtering days. The
30
31
32 99 collected samples were immediately frozen in liquid nitrogen and stored at -80°C in a deep
33
34 100 freezer until RNA extraction. For the RNA-seq analysis we selected the animals according to
35
36 101 the estimated breeding value (EBV) for backfat thickness (BFT) calculated by ANAS as
37
38 102 described by Russo et al. (2000; 2008). EBVs were determined through a BLUP multiple-trait
39
40
41 103 animal model procedure (Henderson & Quaas, 1976) using the BFT, measured in mm,
42
43 104 recorded post mortem in correspondence of the *gluteus medius* muscle. The model included
44
45 105 fixed effects of batch in test, sex, age at beginning of test, age of sow, weight at slaughter,
46
47 106 age at slaughter, and inbreeding coefficient as well as the random effects of litter, individual
48
49
50 107 permanent environment, and animal. Pigs’ genetic merit for the BFT trait was calculated
51
52 108 taking into account the additive relationship matrix. EBVs were expressed as differences
53
54
55 109 from the genetic mean value for the considered trait in the year 1993. Backfat thickness
56
57 110 genetic index may present negative values because the value of the trait is referred to the
58
59
60

1
2
3 111 fixed genetic base defined by ANAS as mean values of the pigs born in 1993 and considered
4
5 112 as “zero”, so the more negative values indicate lower values of BFT. The animals were
6
7
8 113 selected to compose two groups of 10 pigs showing extreme and divergent characteristics
9
10 114 for the BFT EBV with respect to the larger population of the 949 pigs (Table 1). The twenty
11
12 115 animals considered for RNA-seq analysis were slaughtered in 12 dates, with 5 dates
13
14 116 common to both groups. The animals were selected also according to their pedigree, in
15
16 117 order to avoid the presence of full sibs in the considered groups. From now on the two
17
18 118 groups will be referred as FAT and LEAN samples.
19
20
21
22 119
23
24 120 RNA extraction, library preparation, sequencing
25
26
27
28 121 Total RNA was extracted with Trizol (Invitrogen) according to the manufacturer’s
29
30 122 instruction. RNA extracted samples were quantified using a Nanodrop ND-1000
31
32 123 spectrophotometer (Nanodrop Technologies) and the quality of the RNA was assayed using
33
34 124 an Agilent 2100 BioAnalyzer (Agilent Technologies). The RNA libraries were prepared from
35
36 125 total RNA using the TruSeq RNA sample preparation kit (Illumina) and version 3 of the
37
38 126 reagents, following the manufacturer’s suggested protocol. The libraries were tagged and
39
40 127 couples of libraries were run on a single lane of an Illumina HiSeq2000. Reads are 100 nt
41
42 128 paired-end represented in FASTQ format.
43
44
45
46 129
47
48
49 130 Architecture of the bioinformatics pipeline
50
51
52
53 131 A computational pipeline to process the sequencing data for gene/transcript expression
54
55 132 estimation and to perform differential expression analysis between the two sample groups
56
57 133 was developed. The pipeline components to achieve expression estimates were assembled
58
59
60

1
2
3 134 using Scons software (<http://www.scons.org/>. Accessed 22 June 2015), which allows the
4
5 135 parallelization and automation of the pipeline tasks. The pipeline and its following steps are
6
7 136 detailed in the next paragraphs.
8
9
10 137
11
12 138 RNA-seq data pre-processing and mapping to swine genome
13
14
15
16 139 Exploratory analyses on the raw reads quality were carried out using the FastQC v0.10.1
17
18 140 software (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 22 June
19
20 141 2015), which generates an HTML report for each sample read set. Read fragments with
21
22 142 quality Phred score lower than 30 were trimmed using the DynamicTrim script of the
23
24 143 SolexaQA v2.1 (Cox et al., 2010). The FASTX-Toolkit v0.0.13.2
25
26 144 (http://hannonlab.cshl.edu/fastx_toolkit/. Accessed 22 June 2015) was used for trimming
27
28 145 result report. A custom Python script using the HTSeq package (Anders et al., 2015) filtered
29
30 146 out the trimmed reads shorter than 50 nucleotides. To maintain a consistent paired-end
31
32 147 read set, discarded read mates were also filtered out, despite their length and quality. Each
33
34 148 sample paired-end clean read set was mapped to the swine genome (Sscrofa10.2.70) by
35
36 149 Tophat v2.0.8 (Kim et al., 2013) using default parameters with transcriptome inference from
37
38 150 Ensembl annotation (Tophat2 used Bowtie v2.1.0.0; Langmead & Salzberg, 2012) and
39
40 151 SAMtools v0.1.19(Li et al., 2009).
41
42 152
43
44 153 Gene/transcript expression evaluation and transcript reconstruction
45
46
47
48
49 154 Gene annotation for the reference genome was retrieved from Ensembl (BioMart) using the
50
51 155 biomaRt R package (Durinck et al., 2009). Read alignments were processed by Cufflinks
52
53 156 v2.1.1 (Roberts et al., 2011a; Roberts et al., 2011b; Trapnell et al., 2010) to identify and
54
55
56
57
58
59
60

157 discover expressed genes and transcripts, and to quantify their expression. Expression data
158 were indicated as Fragments Per Kilobase of transcript per Million mapped reads (FPKM).
159 Cufflinks was applied to each sample alignment; then, we merged the transcript predictions
160 in a non-redundant reference using the Cuffmerge tool from the Cufflinks package. To
161 reduce artefacts deriving from the transcript prediction and normalisation strategies, only
162 predicted transcripts at least 200 nt long and with minimal expression of 100 (Cufflinks
163 normalised) reads in at least one of the two groups were considered for transcriptome
164 reconstruction and for the following analyses.

165

166 Gene and transcript differential expression assessment

167 The samples were inspected by principal component analysis to examine their similarities.
168 The read counts of each gene in the 20 considered samples were transformed with the
169 variance stabilizing transformation function provided by the DESeq2 package (Anders &
170 Huber, 2010) and used to compute the principal components.
171 The genes identified by Cufflinks were assessed for differential expression (DE) between the
172 LEAN and FAT groups, by means of two strategies, namely Cuffdiff2 (v2.1.1 from the
173 Cufflinks package; Trapnell et al., 2012) and DESeq2 v1.2.1 (Anders & Huber, 2010). Instead,
174 transcript DE was assayed only with Cuffdiff2. To represent gene expression, the two
175 methods use similar statistical approaches based on generalized linear model (GLM) of the
176 negative binomial family. Cuffdiff2 extends the model using a beta negative binomial
177 distribution to handle uncertainty of multi-mapped reads. On the contrary, DESeq2
178 considers only uniquely mapped reads (counted by means of the htseq-count script of the
179 HTSeq package (Anders et al., 2015), but facilitate the specification in the statistical model

1
2
3 180 of additional factors effecting the fit of the GLM. In this study, the statistical model included
4
5 181 sex effect as a potential conditioning factor. Gene and transcript DE test computed P-values
6
7
8 182 were corrected according to the Benjamini-Hochberg procedure. Differentially expressed
9
10 183 genes and transcripts were considered statistically significant according to false discovery
11
12 184 rate less than or equal to 0.05.
13

14 185
15
16
17 186 Transcript characterisation
18

19
20 187 Using custom scripts including BEDTools v2.17.0 software (Quinlan & Hall, 2010), we
21
22 188 retrieved the nucleotide sequences of the transcripts extracting from the *Sus scrofa* genome
23
24 189 the stretches of nucleotides according to the annotation generated by the RNA-seq analysis
25
26 190 tools. Transcripts were identified or characterised by sequence similarity using BLASTN and
27
28 191 BLAST2 from the NCBI BLASTN suite
29
30 192 (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_L
31
32 193 OC=blasthome. Accessed 22 June 2015) using Megablast algorithm (Morgulis et al., 2008).
33
34 194 To assign a gene name, the sequences IDs obtained with this comparison were used to
35
36 195 query the NCBI Gene and the UniGene databases (<http://www.ncbi.nlm.nih.gov/unigene/>.
37
38 196 Accessed 22 June 2015). We used two strategies for transcript annotation. DE transcripts
39
40 197 and genes were annotated by similarity using nr/nt nucleotide collection. The threshold
41
42 198 considered for the identification of our transcripts was identity $\geq 80\%$ in at least 70% of the
43
44 199 sequence length of a transcript present in the database. Transcripts from new genes were
45
46 200 characterized using a comparative genomics approach. We compared the new transcripts
47
48 201 from intergenic regions with known human transcripts (RefSeq Release 72) by aligning with
49
50 202 BLASTN (NCBI BLAST 2.2.29+). For each transcript the best hit was considered, and then
51
52
53
54
55
56
57
58
59
60

alignments with E-value greater than 10e-6, identity less than 60%, and length less than 100 nucleotides were discarded.

205

206 Prediction of coding/non-coding potential

207 The transcript coding potential was predicted by CPC (Coding Potential Calculator; Kong et al., 2007). CPC is a support vector machine-based classifier of transcript protein-coding potential grounding on six features of sequence. Three features assess the extent and quality of the predicted transcript ORF: the Framefinder software identifies the longest ORF in the three forward and in the three reverse frames, then the coverage and the integrity of the predicted ORF are evaluated. Another three features derive from results of BLASTX search against UniProt Reference Clusters. All the features contribute together to a final score, and to the classification of transcripts as coding or non-coding. Only transcripts not including uncalled bases were considered for CPC analysis.

216

217 Validation by quantitative real time-PCR

218 The validation of selected RNA-seq results was performed using a quantitative real time-PCR (qRT-PCR) approach using 18 out of the 20 samples used for the RNA-seq analysis. Two samples, one in the FAT group and one in the LEAN group, were not considered because the total RNA extracted was used completely for the RNA-seq analysis. QRT-PCR validation was carried out using Rotor-Gene TM 6000 (Qiagen - Corbett Research). After DNase treatment (TURBO DNA-free™, Ambion, Applied Biosystems), 1 µg of total RNA was reverse transcribed using the iScript cDNA Synthesis kit (BIORAD) according to the manufacturers' instructions.

1
2
3 226 The samples were first used to analyze four candidate normalizing genes beta-2-
4
5 227 microglobulin (*B2M*), polymerase (RNA) II (DNA directed) polypeptide A, 220kDa (*POLR2A*),
6
7 228 hypoxanthine phosphoribosyltransferase 1 (*HPRT1*), tyrosine 3-monooxygenase/tryptophan
8
9 229 5-monooxygenase activation protein, zeta (*YWHAZ*). The primer pairs and the PCR
10
11 230 conditions used are reported in Supplementary Table 1. The expression levels of these four
12
13 231 genes were evaluated using NormFinder and *B2M* and *HPRT1*, the two most stably
14
15 232 expressed normalizing genes, were utilized as reference genes. For each gene selected for
16
17 233 validation, we designed an external primer pair to obtain the amplicon for the standard
18
19 234 curve construction and an internal primer pair for the qRT-PCR on Rotor Gene 6000 (Table
20
21 235 S1). Standard curves for each gene were generated from 10-12 serial dilutions (from 10⁹ to
22
23 236 25 molecules/μl) of the PCR amplicons obtained with the external primer pairs and
24
25 237 containing the internal primers used in the qRT-PCR analysis. Amplifications were performed
26
27 238 in a total volume of 10 μl containing using 5 μl of the SYBR® Premix Ex Taq™ (Takara Bio
28
29 239 Inc.), 0.5 μl of each primer and about 100 ng of cDNA. The used Premix Ex Taq™ is optimized
30
31 240 for a two-step cycling, and the amplification conditions for the tested genes are reported in
32
33 241 Table S1. The PCR efficiency was calculated as $E=10 \exp(-1/\text{slope})$, with a range between -
34
35 242 2.7 and -4.3, indicating a good PCR efficiency result. All the PCR products were checked on a
36
37 243 polyacrylamide gel and the specificity of the amplification was checked by a final melting
38
39 244 curve analysis.
40
41 245 Threshold cycles obtained for the samples were converted by Rotor Gene 6000 to mRNA
42
43 246 molecules/μl using for each gene the relative standard curve (Bustin & Nolan, 2004).
44
45 247 Moreover, the average mRNA molecules/μl for each sample was normalized dividing the
46
47 248 mRNA molecules of a gene /μl by the geometric average of *B2M* and *HPRT1* mRNA
48
49 249 molecules/μl in the given sample, as suggested by Bustin & Nolan, 2004 and Vandesompele
50
51
52
53
54
55
56
57
58
59
60

et al., 2002. Differences on the expression level calculated for FAT and LEAN samples were tested by two-tailed Student's t test. Statistical analyses were performed with SAS version 9.3 (SAS 9.3 Help and Documentation, Cary, NC. SAS Institute Inc.) and nominal P-value ≤ 0.05 was considered as significance threshold.

254

255 Functional characterization

Functional annotation, classification and annotation clustering of selected gene sets were carried out by DAVID Tools 6.7 (Huang et al., 2009) using Biological Processes, Molecular Function gene ontology categories and KEGG pathways. A threshold for significance of $P < 0.01$ and $P < 0.05$ after Benjamini correction was considered for the selection of the functional categories respectively in the characterization of most expressed transcripts and for the selection of the functional categories of DEG.

262

263 Results

264 Samples

In this study we applied RNA-seq by Illumina technology to the study of gene expression in backfat tissue of 20 ILW pigs. We considered a large group of 949 sampled animals, with EBV for BFT ranging from -10.64 mm to 7.28 mm, with mean value and standard deviation (SD) -1.96 mm and 3.01, respectively. We selected, from the whole collected population, two groups of 10 unrelated pigs (FAT and LEAN) with extremely divergent EBVs for BFT, with 1:1 sex ratio within each group. The mean values of each of the two selected groups of pigs are outside the range -7.98 mm / 4.06 mm defined by the mean value of the 949 samples ± 2

1
2
3 272 standard deviations. Specifically, FAT and LEAN animals were associated to average BFT
4
5 273 values of +5.22 mm (± 1.30 SD) and -8.63 mm (± 1.40 SD) as indicated in Table 1.
6
7 274
8
9
10 275 Sequencing, reads pre-processing and mapping
11
12
13 276 Pairs of samples were run together, after barcoding, on a single lane of an Illumina HiSeq
14
15 277 2000 apparatus, obtaining a total of 3,917,123,414 raw reads for the 20 considered samples,
16
17 278 with an average of 195,856,171 raw reads per sample (Table S2; GEO accession GSE68007).
18
19 279 After trimming and length filtering the clean reads per sample were on the average
20
21 280 113,934,264 (58.04%) and were used for read-to-genome mapping (Figure S1A). Reads that
22
23 281 align on a single genome locus (uniquely mapped reads) were on the average the 91.07% of
24
25 282 the mapped reads (Table S2). The 72.42% of the uniquely mapped reads (72,219,306.45 on
26
27 283 the average aligned to annotated exons, the 19.15% mapped on intergenic regions and the
28
29 284 8.43% mapped on introns of annotated genes. The deep sequencing allowed the
30
31 285 identification of genes expressed at low level and relatively rare alternatively spliced
32
33 286 transcripts. We observed splicing events in the 21.19% of the reads on the average,
34
35 287 providing useful information for the reconstruction of alternative transcript isoforms (Figure
36
37 288 S1B).
38
39 289
40
41
42
43
44
45
46 290 Transcripts and genes expressed in backfat samples
47
48
49
50 291 The deep sequencing analysis of backfat transcripts performed on two groups of pigs
51
52 292 divergent for fat deposition in this tissue allowed the detection of 63,418 transcripts. Many
53
54 293 of them have not yet been annotated in the porcine genome, thus providing new consistent
55
56 294 resources for pig genome annotation and studies of adipose tissue biology. We identified
57
58
59
60

the expression of genes on all porcine autosomes, sex chromosomes and mitochondrial genome. Chromosome 1 has the largest number of expressed genes (8.23%), followed by chromosomes 6 (7.84%) and 2 (7.25%). Furthermore, a non-negligible part (12.48%) of the expressed genes is located in genomic scaffolds (Figure S1C), as about the 7.5% of the genome has no assigned location yet, as described in Ensembl annotation of pig genome (database version 78 at the time of the analysis ; http://www.ensembl.org/Sus_scrofa/Location/Genome. Accessed 22 June 2015). In term of genes, we identified 23,483 expressed pig genes: 12,707 known and 10,776 putative new genes.

Transcripts were split in different classes according to their matching with the genome annotations (Figure 1A, Table S3). Transcripts matching exactly the reference annotation are indicated as “known” transcripts; annotated transcripts’ new isoforms or overlapping with annotated transcript are indicated as “novel isoforms; and all other transcripts, such as those expressed from extragenic regions, are referred as “new” transcripts and might represent putative new genes. The majority of expressed transcripts are novel isoforms (35,030; the 55.2%) or known transcripts (12,969, representing the 20.5%) that are prevalently annotated as protein coding (12,883; 99.3%). The expressed new transcripts are 15,419 (24.3%).

Transcript lengths range from 200 to 50,610 nt, with median and average values of 3,224 and 3,979. Average size exceeds the 2 kb pig mean transcript size that can be estimated according to Ensembl pig coding transcript annotation. We observed that the novel isoforms reconstructed are longer than “known” pig transcripts (Figure 1B).

1
2
3 317 Sequences longer than 5 kb compose the 25% of the expressed transcripts. Noteworthy, we
4
5 318 detected two transcripts overlapping *ZBTB16* gene and two new transcripts from
6
7 319 chromosome 16 that are longer than 40 kb.
8
9
10 320 Considering transcripts expression, we observed that new transcripts are less expressed in
11
12 321 fat tissue than known transcripts (Figure 1C). Nevertheless, all the three transcript
13
14 322 categories span a considerably large range of expression values.
15
16
17 323 The majority of the expressed genes (12,138; 52%) present only one transcript isoform
18
19 324 expressed in fat tissue (Figure 1D); the 27.0% and the 18.3% of the genes present two and
20
21 325 three expressed isoforms, respectively, whereas the remaining 12.7% of the genes are
22
23 326 associated each one to 4 to 31 different isoforms. We identified 31 isoforms for the gene
24
25 327 *MAP4K4*, for which a complex expression pattern is reported in humans: Ensembl release 79
26
27 328 lists 20 *MAP4K4* transcripts generated by at least 3 different promoters, by complex
28
29 329 alternative splicing and by polyadenylation patterns, whereas five protein isoforms are
30
31 330 reported in UniProt release 2015_3.
32
33
34 331 Looking at isoform types, Figure 1E shows that many genes expressing only one transcript
35
36 332 (first bar from the left) in fat tissue are putative new genes (green portion). Interestingly,
37
38 333 some genes expressing only one transcript in fat tissue are represented only by a novel
39
40 334 isoform (first bar, red shadows). The proportion of novel isoforms (red portion) increases
41
42 335 along with the numbers of expressed transcripts per gene. Moreover, the transcripts classes
43
44 336 showing exonic overlap compared to a reference transcript are found in genes with a
45
46 337 varying number of transcripts and are particularly abundant in genes with up to three
47
48 338 isoforms. The remaining transcript classes are very rare.
49
50
51 339 Interesting new isoforms derived from known genes regard Perilipin 2 (*PLIN2*; alias *ADFP*,
52
53 340 adipofilin), an important gene for fat metabolism in pigs (Davoli et al., 2010; Gandolfi et al.,
54
55
56
57
58
59
60

2011) whose expression in humans correlates positively with cytosolic triacylglycerol levels (Conte et al., 2013). Only one transcript is currently annotated in Ensembl for pig *PLIN2* (ENSSSCT00000005701), whereas according to our results, *PLIN2* expressed four different isoforms. The most expressed *PLIN2* transcript (expressed two times more in FAT than in LEAN pigs) is a non-annotated isoform (TCONS_00002441 in Table 2; 2441DE in Figure S3) characterized by the skipping of the fourth exon. The same transcript has also a shorter 3' sequence with respect to the canonical *PLIN2/ADFP* form, probably due to the use of an alternative polyadenylation site. Importantly, the skipping of the 83 nt long exon four introduces downstream a shift in the reading frame and a premature stop codon. Thus, this transcript encodes a truncated protein (only 80 aa) corresponding to the N-terminal region and of the Perilipin domain of the *PLIN2* protein annotated isoform (463 aa). The other two new transcripts differ from the annotated isoform, one for the skipping of exon 2, and the other for a longer first exon, probably due to alternative TSS usage by different promoters. The four expressed isoforms are also heterogeneous in the length of the 3' UTR region.

355

356 Coding and non-coding transcripts from new genes

357 We obtained a characterization of intergenic transcripts from new genes first both by
358 similarity, comparing them against human transcripts, and by predicting their coding
359 potential. New pig transcripts with an assigned human best hit were 10,020 (65%),
360 expressed by 7,099 genes (66%), and corresponding to 4633 human Refseq sequences
361 (3,882 unique gene symbols; Table S4).

362 We considered 12,702 intergenic transcripts for protein coding potential analysis. For each
363 transcript, the coding potential of both the forward and the reverse complement sequence

1
2
3 364 were evaluated. According to CPC results, we classified the 35.8% (4,551) of transcripts as
4
5 365 coding, and the 64.2% (8,151) as non-coding. As done by Zhou et al., (2014), we considered
6
7
8 366 as proper non-coding only those transcripts classified as non-coding and having a CPC score
9
10 367 lower than -1 for both the forward and the reverse sequence. A portion of the non-coding
11
12 368 transcripts (37.5%) resulted with CPC score < -1 for both the forward and the reverse
13
14
15 369 complement sequences. We refer to these transcripts as “reliable non-coding” class, which
16
17 370 represented 24% (3,056) of the intergenic transcripts (Figure 2A). We observed that
18
19
20 371 intergenic coding transcripts are on average longer than intergenic non-coding transcripts
21
22 372 (4,149 and 3,083 nt, respectively), and that the reliable non-coding fraction has a even
23
24 373 shorter average length (2,571 nt; Figure 2B and Table S5). Reportedly, non-coding
25
26 374 transcripts tend to be shorter and to have fewer exons than coding transcripts in
27
28
29 375 mammalian genomes (Iyer et al., 2015).
30
31 376 Coding transcripts have an average expression in fat tissue higher than the non-coding
32
33 377 transcripts (5.32 and 2.28 FPKM respectively, and 3.23 FPKM for the reliable non-coding
34
35 378 group; Figure 2C). One reliable non-coding transcript is ranked within the 100 most
36
37 379 expressed transcripts detected in backfat tissue; 15 reliable non-coding transcripts are
38
39
40 380 within the 1,000 most expressed transcripts; and 98 are within the 10% most expressed
41
42
43 381 transcripts (Table S6). In agreement with previous results showing that coding transcripts
44
45 382 tend to present higher expression than non-coding ones (Cabili et al., 2011; Iyer et al.,
46
47 383 2015), we observe that intergenic transcripts ranking in the 10% most expressed in backfat
48
49
50 384 tissue are enriched in the coding category (55%) and particularly if compared with the
51
52 385 proportion of the coding category within the set of intergenic transcripts (35.8%; Figure 2D,
53
54 386 green portions).
55
56
57 387
58
59
60

388 Function of most expressed transcripts

389 A global view of the transcription profile of porcine backfat tissue was obtained by
390 averaging the FPKM values of all 20 analysed samples. The 1411 most expressed transcripts,
391 accounting together for 75% of expression, were chosen to extract the most expressed
392 genes (Table S6).

393 Among these genes, 59 are indicated as reliable non-coding (CPC score <1) and 66 showing a
394 positive CPC score are indicated as putative coding.

395 According to DAVID functional annotation and clustering, we characterized the biological
396 processes (Table S7) associated to the most expressed genes. Ribosomal activity, oxidative
397 phosphorylation, protein metabolic processes, intracellular protein transport, regulation of
398 translation initiation, fatty acid metabolism, response to oxidative stress resulted to be the
399 biological processes more represented in subcutaneous adipose tissue of the analysed
400 samples.

401

402 Gene/transcript differential expression

403 Unsupervised analysis of gene expression profiles was carried out to inspect similarities
404 among the samples. Principal component analysis revealed a clear separation of the LEAN
405 and FAT samples according to the first two most informative components (Figure S4 A),
406 which, notably, do not separate the samples by sex (Figure S4 B).

407 Average gene expression values for FAT and LEAN groups were 32.46 and 33.63 FPKM). In
408 both groups, few highly expressed genes contribute to the majority of the cumulative
409 expression. For instance, roughly 25% of expressed genes (5,908 and 5,728 in FAT and LEAN,
410 respectively) constitute 95% of the total detected expression (Figure S2). As expected,

1
2
3 411 transcript expression distribution is similar to the gene expression distribution being
4
5 412 positively skewed, with mean and median corresponding to 11.84 and 0.64 FPKM,
6
7 413 respectively. Transcripts average expression values are lower than genes expression values
8
9 414 since the latter was computed as the sum of transcripts expression of each gene.
10
11 415 To identify a set of robust DEG and DET the transcription profiles of FAT and LEAN samples
12
13 416 were compared with the integration of two methods applied at gene and at transcript
14
15 417 levels. Cuffdiff2 identified 414 DEGs between FAT and LEAN groups, corresponding to 1,187
16
17 418 transcripts: 266 DEGs are more expressed and 148 DEGs are less expressed in FAT samples.
18
19 419 Fold changes in base two logarithmic scale of DEGs range from 0.46 to 8.95 for the higher
20
21 420 expressed genes, and from -6.19 to -0.47 for the less expressed ones (Table S8). DESeq2
22
23 421 identified 586 DEGs (185 in common with the DEGs identified by Cuffdiff2) corresponding to
24
25 422 1,504 transcripts: 358 genes are up-regulated and 228 genes are less expressed in FAT
26
27 423 samples. DEGs base two logarithmic scale transformed fold changes (Log_2 FC) range from -
28
29 424 1.13 to -0.20 for the less expressed genes and from 0.21 to 1.18 for the higher expressed
30
31 425 genes (Table S9). Cuffdiff2 differential expression analysis at the transcript-level identified
32
33 426 154 DE transcripts (corresponding to 153 genes): 48 were less expressed and 106 transcripts
34
35 427 were more expressed in FAT samples, with Log_2 FC ranging from -3.44 to -0.54 and from
36
37 428 0.64 to 3.66, respectively (Table S10). On the whole, 818 genes were DE, or associated to at
38
39 429 least one DE transcript, according to at least one method, were detected (Figure 3A).
40
41 430 The overlapping of the different lists of DEGs and the list of DE transcripts (DET) evidenced a
42
43 431 group of 86 DET that are identified by all the approaches, from now on referred as
44
45 432 “common DET” (cDET). These DET belongs to 78 DEG, from now on referred as “common
46
47 433 DEG” (cDEG) since five genes are represented by more than one isoform (Table 2).
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 434 The cDET present the same fold change sign of the corresponding cDEG (Figure 3B): 72 DET
4
5 435 were more expressed in FAT (max Cuffdiff2 gene-level Log₂ FC 2.55 for *DSC2* gene) and 14
6
7
8 436 DET were less expressed in FAT (minimum Log₂ FC -3.44 for an intergenic gene located in
9
10 437 GL894890.1 scaffold). Among the 86 cDET, 44 are known transcripts, 16 are novel isoforms
11
12 438 and 26 come from intergenic regions.

13
14
15 439 cDEG are found in all chromosomes except for chromosomes 16 and Y, with up to 11 DE
16
17 440 genes in chromosome 4 and 19 DE genes in scaffolds (Figure 3C). The most expressed
18
19 441 (average FPKM greater than 100) known cDEG, reported in decreasing FPKM order, are
20
21 442 *DNAJB1*, *CTSH*, *CTGF*, *C1QC*, *SPP1* and *CDO1*.

22
23
24 443

25 26 27 444 Coding and con-coding intergenic DET

28
29
30 445 We considered the 41 novel isoforms or new transcript cDET for CPC analysis. In 14 of these
31
32 446 transcripts both the forward and reverse sequence is probably non-coding, according to
33
34 447 integrated ORF analyses and to similarity searches, and to CPC score thresholds used before.
35
36
37 448 Five cDET with CPC score <-1 were scored as “reliable non-coding”. Of the remaining
38
39 449 transcripts, nine presented low coding potential both in the forward and in the reverse
40
41 450 complement sequence but with CPC score ranging from -1 to 0 (“non-coding”), and 27 were
42
43 451 classified as coding transcripts (Table S11).

44
45
46 452

47 48 49 453 qRT-PCR confirmation of DE for selected genes

50
51
52 454 To validate the results obtained by RNA-seq, eleven cDEG were chosen according to the
53
54 455 absolute value of the Log₂ FC between FAT and LEAN pigs, or for their functional role and
55
56
57 456 involvement in relevant pathways. As reported in Figure 4, the DE of all selected genes has
58
59
60

1
2
3 457 been validated, with high correlation between the fold changes obtained by RNA-seq and by
4
5 458 qRT-PCR data.
6
7 459
8
9
10 460 DE transcript characterisation
11
12
13 461 We characterized the cDEG in terms of their functional role in adipose tissue. Using DAVID
14
15 462 Bioinformatics Resources we first identified the functional categories, enriched in genes
16
17 463 differentially regulated between FAT and LEAN groups.
18
19
20 464 The Biological Process categories enriched in higher expressed DEG are response to
21
22 465 stimulus, immune system process and cell activation, skeletal system development (Table
23
24 466 3). DAVID clustering of the few lower expressed genes detected (*ADSSL1*, *CDO1*, *DNAJB1*,
25
26 467 *HSPA1A*, *HSPA1B*, *HSPA2*, *HSPB8*, *IGFBP5*, *OLFML3*) allowed to identify the functional
27
28 468 categories unfolded protein binding and stress response represented by five heat shock
29
30 469 protein genes that are involved in protein stabilization after cellular stress.
31
32
33 470 Apart from the Gene Ontology-based functional characterization of the whole subsets of
34
35 471 higher- and lower-expressed genes we considered cDEG function and involvement in
36
37 472 specific pathways, according to literature and knowledge bases.
38
39
40 473 Several more expressed genes in FAT animals (*ACP5*, *BCL2A1*, *CD1A*, *EGR2*, *ENPP1*, *GPNMB*,
41
42 474 *INHBB*, *LYZ*, *MSR1*, *OLR1*, *PIK3AP1*, *PLIN2*, *SPP1*, *STC1*) are characterized by a metabolic
43
44 475 function mainly related to adipocyte growth regulation, while others (*CCR1*, *CD163*,
45
46 476 *SLC11A1*) are known to be involved in immune defence of the organism.
47
48
49
50
51 477

52
53
54 478 **Discussion**
55
56

57
58 479 Transcriptome data highlight the adipose tissue complexity
59
60

1
2
3 480 The deep sequencing analysis of pig backfat transcriptome performed allowed finding
4
5 481 thousands of genes and transcripts expressed. In the present study, we applied stringent
6
7 482 cleaning and filtering procedures of the sequencing data and, on average, 90 million reads
8
9 483 per sample were mapped, obtaining a higher sequencing depth compared to previous
10
11 484 studies (Chen et al., 2011; Jiang et al., 2013; Sodhi et al., 2014; Wang et al., 2014).
12
13 485 The adipose tissue is not only metabolically and transcriptionally active, but has been
14
15 486 recognized as an important endocrine organ (Kershaw et al., 2004; Trayhurn et al., 2005).
16
17 487 Adipocytes are a dynamic and highly regulated population of cells (Rosen & MacDougald,
18
19 488 2006; Moreno-Navarrete & Fernández-Real, 2012). Our results agree with these data
20
21 489 supporting the characterization of the adipocytes as highly specialised endocrine cells that
22
23 490 can play key roles in various physiological processes. The multifunctionality and the
24
25 491 complexity of the tissue is witnessed also by the high number of transcripts (more than sixty
26
27 492 thousands) found in the present study, including many new transcripts from previously non-
28
29 493 annotated loci in porcine genome. The majority of the reconstructed sequences are novel
30
31 494 isoforms of already known genes that express more than two different transcripts each.
32
33 495 Similar patterns observed in human cells (Djebali et al., 2012) and the high quality of the
34
35 496 sequenced reads used in our analysis support the idea that this is more attributable to an
36
37 497 incomplete annotation of the transcript isoforms expressed in pig backfat, than to transcript
38
39 498 reconstruction artefacts. The different isoforms derived from the same locus arisen from
40
41 499 our analysis and observed for almost half of the expressed genes, may contribute to
42
43 500 improve the knowledge of the porcine transcriptome, and to refine the current swine
44
45 501 genome annotation. The new *PLIN2* isoforms reported above are an interesting example,
46
47 502 especially if compared to the human genome where at least eight *PLIN2* transcript isoforms
48
49 503 are annotated and only four of them are coding. Remarkably, three human *PLIN2* isoforms
50
51
52
53
54
55
56
57
58
59
60

1
2
3 504 encode N-terminal truncated amino acid chains that are similar to the truncated isoform we
4
5 505 reconstructed in our study, and whose function has not yet been elucidated. Furthermore,
6
7 506 Russell et al. (2008) identified in a *PLIN2* deficient mouse cell line the expression of a *PLIN2*
8
9
10 507 C-terminal truncated protein that may partially replace the function of the full-length
11
12 508 protein. Additional studies are needed to understand if and how the short transcript we
13
14 509 found differentially expressed could change the gene functions compared to the wild type
15
16
17 510 long protein.

18
19
20 511

21
22 512 Functional characterisation of the adipose tissue expression profile
23

24
25 513 The profile of the subcutaneous adipose tissue transcriptome in pigs was delineated and the
26
27 514 functional analysis of the genes expressed in backfat tissue was performed to know their
28
29 515 metabolic role and to connect them to specific competences of the tissue. We didn't find
30
31 516 particular differences between the functional categories of the genes expressed in the
32
33 517 backfat tissue of FAT and LEAN pigs. More in details among the most expressed genes in the
34
35 518 fat tissue, many are involved in metabolic pathways and biological processes related to
36
37 519 protein metabolism, oxidoreductase activity for ATP production, regulation of lipid synthesis
38
39 520 and degradation.

40
41
42 521

43
44
45 522 Genes differentially expressed between LEAN and FAT animals converge and
46
47 523 connect to specific functions
48
49

50
51
52 524 The detection of DE genes and transcripts has been obtained by a stringent procedure
53
54 525 grounding on integration of different methods for expression estimation and differential
55
56 526 expression testing, as done in a recent study (Ropka-Molik et al., 2014) focused to muscle
57
58
59
60

1
2
3 527 tissue gene expression in pigs of different breeds. In the present study, which compares pigs
4
5 528 of the same breed and reared under standard conditions, we detected significant gene
6
7 529 expression variations. The sensitivity of our approach was supported by the successful
8
9 530 validation of all the eleven DEG assayed.
10
11 531 We analyzed the biological functions of genes differentially expressed between FAT and
12
13 532 LEAN animals (Figure 5). It is interesting to note that the main differences were found for
14
15 533 functional categories of genes related Inflammation and immunity that resulted more
16
17 534 expressed in FAT pigs. The genes less expressed in FAT animals include some heat shock
18
19 535 protein genes. The biological functions of DEGs show a stronger activation in adipose tissue
20
21 536 of FAT pigs of genes for important processes involved in hypertrophy and adipogenesis, such
22
23 537 as differentiation and maturation. Supposedly, these biological processes could be altered in
24
25 538 adipose tissue of FAT pigs due to dysregulated adipose metabolism and endocrinology
26
27 539 similarly to what was hypothesized in humans (Sethi, 2010). On the whole, there is a
28
29 540 consistent difference concerning the biological functions characterizing the most expressed
30
31 541 genes on backfat tissue and those of the genes differentially expressed between FAT and
32
33 542 LEAN pigs.
34
35 543
36
37 544 Some genes higher expressed in FAT animals could modulate backfat
38
39 545 physiological processes
40
41
42
43 546 Specific DEGs more expressed in FAT pigs participate to biochemical pathways related to
44
45 547 and involved in adipocytes metabolism and adipose tissue physiology. Ectonucleotide
46
47 548 pyrophosphatase/phosphodiesterase 1 (*ENPP1*) encodes a catalytic enzyme involved in
48
49 549 adipocyte maturation (Liang et al., 2007). Pan et al. (2011) showed that the over-expression
50
51
52
53
54
55
56
57
58
59
60

1
2
3 550 of *ENPP1* in a human cell line resulted in adipocyte insulin resistance and demonstrated an
4
5 551 association with fatty liver, hyperlipidemia, and dysglycemia. Accordingly, the study of
6
7 552 Chandalia et al. (2012) underlined an increased *ENPP1* expression in adipose tissue
8
9 553 associated with defective adipocyte maturation leading to pathogenesis of insulin resistance
10
11 554 and its associated complications for glucose and lipid metabolism in absence of obesity. In
12
13 555 addition, Meyre et al. (2005) reported the presence of three *ENPP1* SNPs in human gene
14
15 556 associated with adult obesity and increased risk of glucose intolerance and type 2 diabetes.
16
17 557 Furthermore, also the genes acid phosphatase 5, tartrate resistant (*ACP5*) and lysozyme
18
19 558 (*LYZ*) that in this research have higher transcriptional level in FAT pigs have been reported to
20
21 559 be involved in excessive backfat deposition in pigs and in the development of
22
23 560 atherosclerosis (Padilla et al., 2013).
24
25
26 561 In the present research, some genes overexpressed in the adipose tissue of FAT pigs, namely
27
28 562 *STC1*, *EGR2*, and *INHBB*, are related to adipocyte differentiation and adipocyte maturation.
29
30 563 *STC1* (Stanniocalcin 1) has been reported in literature to be up-regulated during
31
32 564 adipogenesis and to modulate steroidogenesis. Serlachius & Andersson (2004) related *STC1*
33
34 565 up-regulation to the set of survival genes in adipocyte differentiation, which is also
35
36 566 associated to overexpression of the anti-apoptotic proteins *BCL2* reported to be involved in
37
38 567 inflammation pathway. *EGR2* (early growth response 2) is a direct target of *mir-224-5p*, a
39
40 568 negative regulator of adipocyte differentiation that is down regulated during the early
41
42 569 process of mouse adipocyte differentiation, and the expression of *EGR2* is increased (Peng
43
44 570 et al., 2013). The *INHBB* (Inhibin beta B) gene coding for the activin B subunit is part of the
45
46 571 inhibins/activins family of proteins with cytokine and hormone activity. In human and mice,
47
48 572 *INHBB* has been associated to the physiological and metabolic modifications during
49
50 573 adipogenesis when it is highly expressed and is the predominant activin in human adipose
51
52
53
54
55
56
57
58
59
60

1
2
3 574 tissue (Hoggard et al., 2009). *INHBB* is member of TGF-protein superfamily of secreted
4
5 575 growth factors involved in many biological responses including regulation of apoptosis;
6
7 576 proliferation and differentiation of human adipocytes; tissue remodeling; and inflammatory
8
9
10 577 immune response (Dani C., 2013). It can be hypothesized that in FAT pigs the pro-adipogenic
11
12 578 *INHBB* gene expression increases as it is involved in the differentiation of preadipocytes into
13
14 579 mature adipocyte, and that *INHBB* is involved in many physiological processes and including
15
16
17 580 the control of food intake and to energy metabolism through the regulation of hypothalamic
18
19 581 and pituitary hormone secretions. Another gene overexpressed in FAT pigs related to
20
21 582 feeding and pituitary secretions is *GPNMB* (glycoprotein transmembrane NMB). GPNMB is
22
23 583 one of the receptors activated by bombesin-like endogenous peptide ligands, such as
24
25 584 gastrin-releasing peptide (*GRP*), neuromedin B (*NMB*) and neuromedin C (*GRP18-27*). These
26
27 585 receptors are involved in the regulation of many biological functions including
28
29 586 thermoregulation, feeding, pituitary, gastric and pancreatic secretion. The NMB/NMB-R
30
31 587 pathway is involved in the regulation of a wide variety of behaviours, such as spontaneous
32
33 588 activity, feeding, and anxiety-related behaviour (Yamada et al., 2002).
34
35
36 589 The *OLR1* (Oxidized low density lipoprotein (lectin-like) receptor 1) gene resulted more
37
38 590 expressed in FAT pigs compared to LEAN animals. This gene codes for a LDL receptor that
39
40 591 belongs to the C-type lectin superfamily, one of many target genes, including perilipins, of
41
42 592 the PPAR signalling, which is involved specifically in lipid metabolism and fatty acids
43
44 593 transport. In this way, *OLR1* is a receptor that mediates the recognition, internalization and
45
46 594 degradation of oxidatively modified low-density lipoprotein by vascular endothelial cells.
47
48 595 *OLR1* removes oxidised low-density lipoproteins from the circulation, as part of lipid
49
50 596 metabolism pathways (Mehta et al., 2002).
51
52
53
54
55
56
57 597
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

598 Genes involved in immunity and inflammation are more expressed in FAT
599 animals
600 Some other genes overexpressed in FAT pigs are related to immunity. Inflammatory links
601 between human obesity and metabolic diseases are well known mechanisms based on the
602 recruitment of immune cells into adipose tissue (Kabir et al., 2014). The development of a
603 pre-inflammatory condition in presence of dysregulated excessive adipogenesis is
604 associated with adipose macrophage infiltration and activation. From our study, we can
605 hypothesize a similar process in backfat tissue of FAT pigs where we identified the over
606 expression of the gene macrophage scavenger receptor 1 (*MSR1*), a membrane glycoprotein
607 that in humans is involved in the pathologic deposition of cholesterol in arterial walls during
608 atherogenesis (Haasken et al., 2013). Additionally, the overexpression of secreted
609 phosphoprotein 1 (*SPP1*) in FAT pigs can suggest the hypothesis that this gene is acting as a
610 proinflammatory cytokine that promotes monocyte chemotaxis and cell motility and might
611 link, in pigs like in mice, fat accumulation to the development of insulin resistance by
612 sustaining inflammation and the accumulation of macrophages in adipose tissue (Nomiya
613 et al. 2007). Interestingly, a porcine *SPP1* gene polymorphism was associated to backfat
614 thickness in the Landrace × Jeju (Korea) Black pig F2 population (Han et al., 2012). *SPP1*
615 might play a key role in the pathway that leads to type I immunity enhancing interferon-
616 gamma and interleukin-12 production and suppressing interleukin-10 (Ashkar et al., 2000).
617 Therefore, these data allow hypothesizing *SPP1* as a gene associated, in pigs like to in human,
618 to the link between obesity, adipose tissue inflammation, and insulin resistance. In addition,
619 phosphoinositide-3-kinase adaptor protein 1 (*PIK3AP1*), higher expressed in FAT pigs, is a
620 positive regulator of phosphatidylinositol 3-kinase (*PI3K*) signalling. *PI3K* signalling pathway

1
2
3 621 has a key role in the insulin-dependent regulation of adipocyte metabolism (glucose and
4
5 622 lipid metabolism). Besides, *PI3K* participate in obesity-associated inflammatory cell
6
7
8 623 recruitment (neutrophils and macrophages), as well as in the CNS-dependent neurohumoral
9
10 624 regulation of food intake/energy expenditure (McCurdy & Clemm, 2013; Beretta et al.,
11
12 625 2015).

13
14 626 Other genes found in the present research and related to inflammatory condition of the
15
16
17 627 adipose tissue in FAT pigs are particularly interesting to mention. CD163, member of the
18
19
20 628 scavenger receptor cysteine-rich superfamily (Guo et al., 2014; Smith et al., 2014); solute
21
22 629 carrier family 11 (proton-coupled divalent metal ion transporter), member 1 (*SLC11A1*), a
23
24 630 gene involved in the resistance to *Salmonella* infection (Kommadath et al., 2014) as well as
25
26
27 631 the chemokine (C-C motif) receptor 1 (*CCR1*), that was previously found overexpressed in
28
29 632 obese pigs (Kogelman et al., 2014); BCL2-related protein A1 (*BCL2A1*), a gene found to be
30
31 633 overexpressed in pigs with an high obesity index and that is related to immunity,
32
33 634 inflammatory pathway, and osteoclast differentiation (Kogelman et al., 2014); CD1a
34
35 635 molecule (*CD1A*, indicated as *PCD1A* on the cited paper), a surface antigen involved in
36
37
38 636 immunity was found to be overexpressed in obese pigs by Kogelman et al. (2014). The same
39
40
41 637 Authors highlighted a strong connection between fat deposition on the body (obesity),
42
43 638 immunity and bone development. They also indicated that *CCR1* gene is a strong candidate
44
45 639 regulator of immune response as it is a receptor of pro-inflammatory chemokines in adipose
46
47
48 640 tissue playing a pivotal role in obesity-associated diseases (Kabir et al. 2014; Lumeng &
49
50 641 Saltiel, 2011).

51
52 642
53
54
55 643 Heat shock response , protein folding and repair are impaired in FAT animals
56
57
58
59
60

1
2
3 644 Considering the 14 genes less expressed in FAT animals, direct relationships with lipid
4
5 645 metabolism are not apparent. However, the “unfolded protein binding” function is enriched
6
7
8 646 among these genes, which include five functionally linked heat shock proteins (*DNAJB1*,
9
10 647 *HSPA1A*, *HSPA1B*, *HSPA2* and *HSPB8*). Heat shock proteins are involved in stabilization of
11
12 648 existing proteins against aggregation, mediating the folding of newly translated proteins in
13
14
15 649 the cytosol and in organelles, and also in the ubiquitin-proteasome pathway. *DNAJB1*, a
16
17 650 member of the Hsp40 family, is a molecular chaperon involved in protein folding and
18
19
20 651 protein complex assembly. *DNAJB1*, a member of the Hsp40 family, promotes protein
21
22 652 folding and prevent misfolded protein aggregation, as *HSPB8*, a member of the Hsp20
23
24 653 family, does (Vicario et al., 2014). *DNAJB1* also stimulates the ATPase activity of protein of
25
26
27 654 the Hsp70 family to which other genes less expressed in FAT pigs (*HSPA1A*, *HSPA1B*, and
28
29 655 *HSPA2*) belong, indicating a possible functional link between these four genes. Our results
30
31 656 suggest a general impairment of the protein folding and repair in the fattest animals, in
32
33
34 657 accordance to previous observations of studies carried out on human obesity. Obesity is a
35
36 658 pathological human condition in which a chronically positive energy balance induces in
37
38
39 659 adipocytes, the cells in charge to store the excess of energy in fat depots, a persistent stress
40
41 660 activating in turn defence processes as autophagy or apoptosis.
42
43 661 As reviewed by Newsholme & de Bittencourt (2014), if the heat shock response, a key
44
45 662 component of the physiological response to resolve inflammation, is hampered in adipose
46
47
48 663 tissue, the adipocyte metabolic stress triggers fat cell senescence with reduction of the heat
49
50 664 shock proteins activity. In this condition, the advance of inflammasome mediated secretory
51
52
53 665 activity from adipose to other tissues promotes cellular senescence in many other cells of
54
55 666 the organism, aggravating obesity-dependent chronic inflammation. This mechanism could
56
57 667 have been activated also in the FAT pigs of our experiment (Figure 5) due to a genetic
58
59
60

1
2
3 668 aptitude of the fattest animals toward a higher fat deposition and adiposity similar to
4
5 669 obesity. Indeed, a decrease in the synthesis of the mRNAs of the heat shock proteins and an
6
7
8 670 increase of the expression of many genes related to an inflammatory status and to immune
9
10 671 response is a characteristic of the fattest pigs. Increase of the expression of *INHBB* and *SPP1*
11
12 672 denotes for instance the augmented production of cytokines and the higher expression of
13
14 673 *ENPP1* and *PIK3AP1* may indicate a status of insulin resistance, one of the typical signals
15
16 674 connected with obesity.
17
18
19
20 675

21
22 676 Pig backfat deposition and impaired stress response may activate inflammation
23
24

25 677 Our results agree with recent studies showing that several immune system and anti-
26
27 678 inflammatory processes are activated and play a critical role in the response to fat
28
29 679 accumulation in porcine backfat tissue (Sodhi et al., 2014) and in visceral fat tissue
30
31 680 (Toedebusch et al., 2014; Wang et al., 2014). Wang et al. (2014) and Zhou et al. (2013) used
32
33 681 three female Landrace pigs to identify DEG between subcutaneous, visceral and
34
35 682 intramuscular fat indicating that visceral and intramuscular adipose tissues were mainly
36
37 683 associated with inflammatory features of the tissue and immune response. Our data suggest
38
39 684 that also in backfat a predominant role of immunity processes is related to an increased
40
41 685 adipose tissue deposition.
42
43
44
45

46 686 The results obtained seem to sustain the hypothesis that the high fat accumulation in
47
48 687 adipose tissue of pigs can determine the development of an inflammatory process
49
50 688 producing a cascade of defence and adaptive reactions in the tissue, such as activation of
51
52 689 immune system and mesenchymal cells differentiation in adipocytes.
53
54
55
56
57
58
59
60

1
2
3 690 A deeper knowledge of the metabolic processes involved in fat deposition can be very
4
5 691 important to develop the use the pig as model species to study obesity and related
6
7 692 disorders for humans because of similar anatomy and physiology (Spurlock & Gabler, 2008;
8
9 693 Litten-Brown et al., 2010; Varga et al., 2010) and considering the above described
10
11 694 similarities between pigs and humans.
12
13 695 In order to fully elucidate the complex gene network regulating backfat deposition on pigs,
14
15 696 it will be important to extend the basic knowledge by further coding and non-coding
16
17 697 transcriptome characterization. Additional information would probably come from studying
18
19 698 interactions between the differentially expressed long RNAs identified in the present paper
20
21 699 and the regulatory microRNAs expressed in porcine adipose tissue identified on some of the
22
23 700 same animals (Gaffo et al., 2014).
24
25
26
27
28 701 The results of the present work unlock the opportunity that some of the identified
29
30 702 differentially expressed genes might be used as biomarkers (Ibáñez-Escriche et al., 2014) to
31
32 703 improve carcass fat traits in to look for SNPs regulating their expression to be included in
33
34 704 selection schemes to make more sustainable the pig production chain.
35
36
37
38
39
40

41 705
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

706 **Acknowledgements**

707 We thank Dr. Maurizio Gallo from Associazione Nazionale Allevatori Suini (ANAS) and Dr.
708 Luca Buttazzoni (CREA-PCM) for providing the porcine backfat tissue samples, BMR
709 genomics for performing the RNA-Seq. This work was supported by Progetto “AGER -
710 Agroalimentare e ricerca”: Advanced research in genomics and processing technologies for
711 the Italian heavy pig production – Hepiget (Grant N. 2011- 0279).
712

References

- Anders, S., & Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.* 11, R106.
- Anders, S., Pyl, P.T. & Huber, W. (2015) HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169.
- Ashkar, S., Weber, G.F., Panoutsakopoulou, V., Sanchirico, M.E., Jansson, M., Zawaideh, S., Rittling, S.R., Denhardt, D.T., Glimcher, M.J. & Cantor, H. (2000) Eta-1 (osteopontin): an early component of type-1 (cell-mediated) immunity. *Science* 287, 860–864.
- Beretta, M., Bauer, M. & Hirsch, E. (2015) PI3K signaling in the pathogenesis of obesity: The cause and the cure. *Adv Biol Regul.* 58, 1-15.
- Bosi, P. & Russo, V. (2004) The production of the heavy pig for high quality processed products. *Italian Journal of Animal Science* 3, 309–321.
- Bustin, S.A. & Nolan, T. (2004) Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J Biomol Tech* 15, 155–166.
- Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A. & Rinn, J.L. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25, 1915–1927.
- Čandek-Potokar, M. & Škrlep, M. (2012) Factors in pig production that impact the quality of dry-cured ham: a review. *Animal* 6, 327–338.
- Chandalia, M., Davila, H., Pan, W., Szuszkiewicz, M., Tuvdendorj, D., Livingston, E.H. & Abate, N. (2012) Adipose tissue dysfunction in humans: a potential role for the transmembrane protein ENPP1. *J. Clin. Endocrinol. Metab.* 97, 4663–4672.

1
2
3 735 Chen, C., Ai, H., Ren, J., Li, W., Li, P., Qiao, R., Ouyang, J., Yang, M., Ma, J. & Huang, L. (2011)
4
5 736 A global view of porcine transcriptome in three tissues from a full-sib pair with extreme
6
7 737 phenotypes in growth and fat deposition by paired-end RNA sequencing. BMC Genomics
8
9 738 12, 448.
10
11 739 Chen, C.H., Lin, E.C., Cheng, W.T.K., Sun, H.S., Mersmann, H.J. & Ding, S.T. (2006) Abundantly
12
13 740 expressed genes in pig adipose tissue: an expressed sequence tag approach. J. Anim. Sci
14
15 741 84, 2673–2683.
16
17 742 Conte, M., Vasuri, F., Trisolino, G., Bellavista, E., Santoro, A., Degiovanni, A., Martucci, E.,
18
19 743 D’Errico-Grigioni, A., Caporossi, D., Capri, M., Maier, A.B., Seynnes, O., Barberi, L.,
20
21 744 Musarò, A., Narici, M.V., Franceschi, C. & Salvioli, S. (2013) Increased Plin2 expression in
22
23 745 human skeletal muscle is associated with sarcopenia and muscle weakness. PLoS ONE 8,
24
25 746 e73709.
26
27 747 Corominas, J., Ramayo-Caldas, Y., Puig-Oliveras, A., Estellé, J., Castelló, A., Alves, E., Pena,
28
29 748 R.N., Ballester, M. & Folch, J.M. (2013) Analysis of porcine adipose tissue transcriptome
30
31 749 reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty
32
33 750 acid composition. BMC Genomics 14, 843.
34
35 751 Cox, M.P., Peterson, D.A. & Biggs, P.J. (2010) SolexaQA: At-a-glance quality assessment of
36
37 752 Illumina second-generation sequencing data. BMC Bioinformatics 11, 485.
38
39 753 Dani, C. (2013) Activins in adipogenesis and obesity. Int J Obes (Lond) 37, 163–166.
40
41 754 Davoli, R., Gandolfi, G., Braglia, S., Comella, M., Zambonelli, P., Buttazzoni, L. & Russo, V.
42
43 755 (2010) New SNP of the porcine Perilipin 2 (PLIN2) gene, association with carcass traits
44
45 756 and expression analysis in skeletal muscle. Mol Biol Rep 38, 1575–1583.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 757 Durinck, S., Spellman, P.T., Birney, E. & Huber, W. (2009) Mapping identifiers for the
758 integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc
759 4, 1184–1191.
- 760 Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A.M., Tanzer, A.,
761 Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G.K., Khatun, J., Williams, B.A.,
762 Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R.F., Alioto, T.,
763 Antoshechkin, I., Baer, M.T., Bar, N.S., Batut, P., Bell, K., Bell, I., Chakraborty, S., Chen,
764 X., Chrast, J., Curado, J., Derrien, T., Drenkow, J., Dumais, E., Dumais, J., Duttagupta, R.,
765 Falconnet, E., Fastuca, M., Fejes-Toth, K., Ferreira, P., Foissac, S., Fullwood, M.J., Gao, H.,
766 Gonzalez, D., Gordon, A., Gunawardena, H., Howald, C., Jha, S., Johnson, R., Kapranov,
767 P., King, B., Kingswood, C., Luo, O.J., Park, E., Persaud, K., Preall, J.B., Ribeca, P., Risk, B.,
768 Robyr, D., Sammeth, M., Schaffer, L., See, L.-H., Shahab, A., Skancke, J., Suzuki, A.M.,
769 Takahashi, H., Tilgner, H., Trout, D., Walters, N., Wang, H., Wrobel, J., Yu, Y., Ruan, X.,
770 Hayashizaki, Y., Harrow, J., Gerstein, M., Hubbard, T., Reymond, A., Antonarakis, S.E.,
771 Hannon, G., Giddings, M.C., Ruan, Y., Wold, B., Carninci, P., Guigó, R., Gingeras, T.R.,
772 (2012) Landscape of transcription in human cells. Nature 489, 101–108.
- 773 Ferraz, A.L.J., Ojeda, A., López-Béjar, M., Fernandes, L.T., Castelló, A., Folch, J.M. & Pérez-
774 Enciso, M. (2008) Transcriptome architecture across tissues in the pig. BMC Genomics 9,
775 173.
- 776 Gaffo, E., Zambonelli, P., Bisognin, A., Bortoluzzi, S. & Davoli, R. (2014) miRNome of Italian
777 Large White pig subcutaneous fat tissue: new miRNAs, isomiRs and moRNAs. Anim.
778 Genet. 45, 685–698.
- 779 Gandolfi, G., Mazzoni, M., Zambonelli, P., Lalatta-Costerbosa, G., Tronca, A., Russo, V. &
780 Davoli, R. (2011) Perilipin 1 and perilipin 2 protein localization and gene expression study

1
2
3 781 in skeletal muscles of European cross-breed pigs with different intramuscular fat
4
5 782 contents. Meat Sci 88, 631–637.
6
7 783 Gorodkin, J., Cirera, S., Hedegaard, J., Gilchrist, M.J., Panitz, F., Jørgensen, C., Scheibye-
8
9 784 Knudsen, K., Arvin, T., Lumholdt, S., Sawera, M., Green, T., Nielsen, B.J., Havgaard, J.H.,
10
11 785 Rosenkilde, C., Wang, J., Li, H., Li, R., Liu, B., Hu, S., Dong, W., Li, W., Yu, J., Wang, J.,
12
13 786 Stærfeldt, H.-H., Wernersson, R., Madsen, L.B., Thomsen, B., Hornshøj, H., Bujie, Z.,
14
15 787 Wang, X., Wang, X., Bolund, L., Brunak, S., Yang, H., Bendixen, C. & Fredholm, M. (2007)
16
17 788 Porcine transcriptome analysis based on 97 non-normalized cDNA libraries and assembly
18
19 789 of 1,021,891 expressed sequence tags. Genome Biology 8, R45.
20
21 790 Guo, L., Niu, J., Yu, H., Gu, W., Li, R., Luo, X., Huang, M., Tian, Z., Feng, L. & Wang, Y. (2014)
22
23 791 Modulation of CD163 expression by metalloprotease ADAM17 regulates porcine
24
25 792 reproductive and respiratory syndrome virus entry. J. Virol. 88, 10448–10458.
26
27 793 Haasken, S., Auger, J.L., Taylor, J.J., Hobday, P.M., Goudy, B.D., Titcombe, P.J., Mueller, D.L.
28
29 794 & Binstadt, B.A. (2013) Macrophage scavenger receptor 1 (Msr1, SR-A) influences B cell
30
31 795 autoimmunity by regulating soluble autoantigen concentration. J. Immunol. 191, 1055–
32
33 796 1062.
34
35 797 Han, S.-H., Shin, K.-Y., Lee, S.-S., Ko, M.-S., Oh, H.-S. & Cho, I.-C. (2012) Porcine SPP1 gene
36
37 798 polymorphism association with phenotypic traits in the Landrace × Jeju (Korea) Black pig
38
39 799 F2 population. Mol. Biol. Rep. 39, 7705–7709.
40
41 800 Henderson, C.R. & Quaas, R.L. (1976) Multiple Trait Evaluation Using Relatives’ Records.
42
43 801 Journal of Animal Science 43, 1188–1197.
44
45 802 Hoggard, N., Cruickshank, M., Moar, K.M., Barrett, P., Bashir, S. & Miller, J.D.B. (2009)
46
47 803 Inhibin betaB expression in murine adipose tissue and its regulation by leptin, insulin
48
49 804 and dexamethasone. J. Mol. Endocrinol. 43, 171–177.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 805 Hornshøj, H., Conley, L.N., Hedegaard, J., Sørensen, P., Panitz, F. & Bendixen, C. (2007)
4
5 806 Microarray Expression Profiles of 20.000 Genes across 23 Healthy Porcine Tissues. PLoS
6
7 807 ONE 2, e1203.
8
9
10 808 Huang, D.W., Sherman, B.T. & Lempicki, R.A. (2009) Systematic and integrative analysis of
11
12 809 large gene lists using DAVID bioinformatics resources. Nat Protoc 4, 44–57.
13
14 810 Ibáñez-Escriche, N., Forni, S., Noguera, J.L. & Varona, L. (2014) Genomic information in pig
15
16 811 breeding: Science meets industry needs. Livestock Science 166, 94–100.
17
18
19 812 Iyer, M.K., Niknafs, Y.S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., Barrette, T.R., Prensner,
20
21 813 J.R., Evans, J.R., Zhao, S., Poliakov, A., Cao, X., Dhanasekaran, S.M., Wu, Y.-M., Robinson,
22
23 814 D.R., Beer, D.G., Feng, F.Y., Iyer, H.K. & Chinnaiyan, A.M. (2015) The landscape of long
24
25 815 noncoding RNAs in the human transcriptome. Nat. Genet. 47, 199–208.
26
27
28 816 Jiang, S., Wei, H., Song, T., Yang, Y., Peng, J. & Jiang, S. (2013) Transcriptome Comparison
29
30 817 between Porcine Subcutaneous and Intramuscular Stromal Vascular Cells during
31
32 818 Adipogenic Differentiation. PLoS ONE 8, e77094.
33
34
35 819 Kabir, S.M., Lee, E.-S. & Son, D.-S. (2014) Chemokine network during adipogenesis in 3T3-L1
36
37 820 cells: Differential response between growth and proinflammatory factor in
38
39 821 preadipocytes vs. adipocytes. Adipocyte 3, 97–106.
40
41
42 822 Kershaw, E.E. & Flier, J.S. (2004) Adipose tissue as an endocrine organ. J. Clin. Endocrinol.
43
44 823 Metab. 89, 2548–2556.
45
46
47 824 Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. & Salzberg, S.L. (2013) TopHat2:
48
49 825 accurate alignment of transcriptomes in the presence of insertions, deletions and gene
50
51 826 fusions. Genome Biol. 14, R36.
52
53
54 827 Kogelman, L.J.A., Cirera, S., Zhernakova, D.V., Fredholm, M., Franke, L. & Kadarmideen, H.N.
55
56 828 (2014) Identification of co-expression gene networks, regulatory genes and pathways for
57
58
59
60

1
2
3 829 obesity based on adipose tissue RNA Sequencing in a porcine model. BMC Med
4
5 830 Genomics 7, 57.
6
7
8 831 Kommadath, A., Bao, H., Arantes, A.S., Plastow, G.S., Tuggle, C.K., Bearson, S.M.D., Guan,
9
10 832 L.L. & Stothard, P. (2014) Gene co-expression network analysis identifies porcine genes
11
12 833 associated with variation in Salmonella shedding. BMC Genomics 15, 452.
13
14
15 834 Kong, L., Zhang, Y., Ye, Z.-Q., Liu, X.-Q., Zhao, S.-Q., Wei, L. & Gao, G. (2007) CPC: assess the
16
17 835 protein-coding potential of transcripts using sequence features and support vector
18
19 836 machine. Nucleic Acids Res. 35, W345–349.
20
21
22 837 Langmead, B. & Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. Nat.
23
24 838 Methods 9, 357–359.
25
26
27 839 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. &
28
29 840 Durbin, R. (2009) 1000 Genome Project Data Processing Subgroup. The Sequence
30
31 841 Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.
32
33
34 842 Li, X.J., Yang, H., Li, G.X., Zhang, G.H., Cheng, J., Guan, H. & Yang, G.S. (2012) Transcriptome
35
36 843 profile analysis of porcine adipose tissue by high-throughput sequencing. Animal
37
38 844 Genetics 43, 144–152.
39
40
41 845 Liang, J., Fu, M., Ciociola, E., Chandalia, M. & Abate, N. (2007) Role of ENPP1 on Adipocyte
42
43 846 Maturation. PLoS ONE 2, e882.
44
45
46 847 Litten-Brown, J.C., Corson, A.M. & Clarke, L. (2010) Porcine models for the metabolic
47
48 848 syndrome, digestive and bone disorders: a general overview. Animal 4, 899–920.
49
50
51 849 Lumeng, C.N. & Saltiel, A.R. (2011) Inflammatory links between obesity and metabolic
52
53 850 disease. J. Clin. Invest. 121, 2111–2117.
54
55
56 851 McCurdy, C.E. & Klemm, D.J. (2013) Adipose tissue insulin sensitivity and macrophage
57
58 852 recruitment: Does PI3K pick the pathway? Adipocyte 2, 135–142.
59
60

- 1
2
3 853 Mehta, J.L. & Li, D. (2002) Identification, regulation and function of a novel lectin-like
4
5 854 oxidized low-density lipoprotein receptor. *J. Am. Coll. Cardiol.* 39, 1429–1435.
6
7
8 855 Meyre, D., Bouatia-Naji, N., Tounian, A., Samson, C., Lecoœur, C., Vatin, V., Ghoussaini, M.,
9
10 856 Wachter, C., Hercberg, S., Charpentier, G., Patsch, W., Pattou, F., Charles, M.-A.,
11
12 857 Tounian, P., Clément, K., Jouret, B., Weill, J., Maddux, B.A., Goldfine, I.D., Walley, A.,
13
14 858 Boutin, P., Dina, C. & Froguel, P. (2005) Variants of ENPP1 are associated with childhood
15
16 859 and adult obesity and increase the risk of glucose intolerance and type 2 diabetes. *Nat*
17
18 860 *Genet* 37, 863–867.
19
20
21 861 Mikawa, A., Suzuki, H., Suzuki, K., Toki, D., Uenishi, H., Awata, T. & Hamasima, N. (2004)
22
23 862 Characterization of 298 ESTs from porcine back fat tissue and their assignment to the
24
25 863 SSRH radiation hybrid map. *Mamm Genome* 15, 315–322.
26
27
28 864 Moon, J.-K., Kim, K.-S., Kim, J.-J., Choi, B.-H., Cho, B.-W., Kim, T.-H. & Lee, C.-K. (2009)
29
30 865 Differentially expressed transcripts in adipose tissue between Korean native pig and
31
32 866 Yorkshire breeds. *Anim. Genet.* 40, 115–118.
33
34
35 867 Moreno-Navarrete, J.M., & Fernández-Real, J.M. (2012) Adipocyte Differentiation. In
36
37 868 Adipose Tissue Biology (ed. by M.E. Symonds). pp. 17-38. Springer, New York
38
39
40 869 Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T.L., Agarwala, R. & Schäffer, A.A. (2008)
41
42 870 Database indexing for production MegaBLAST searches. *Bioinformatics* 24, 1757–1764.
43
44
45 871 Newsholme, P. & de Bittencourt, P.I.H. (2014) The fat cell senescence hypothesis: a
46
47 872 mechanism responsible for abrogating the resolution of inflammation in chronic disease.
48
49 873 *Current Opinion in Clinical Nutrition and Metabolic Care* 17, 295–305.
50
51
52 874 Padilla, J., Jenkins, N.T., Lee, S., Zhang, H., Cui, J., Zuidema, M.Y., Zhang, C., Hill, M.A.,
53
54 875 Perfield, J.W., Ibdah, J.A., Booth, F.W., Davis, J.W., Laughlin, M.H. & Rector, R.S. (2013)
55
56
57
58
59
60

1
2
3 876 Vascular transcriptional alterations produced by juvenile obesity in Ossabaw swine.
4
5 877 Physiol. Genomics 45, 434–446.
6
7
8 878 Pan, W., Ciociola, E., Saraf, M., Tumurbaatar, B., Tuvdendorj, D., Prasad, S., Chandalia, M. &
9
10 879 Abate, N. (2011) Metabolic consequences of ENPP1 overexpression in adipose tissue.
11
12 880 Am. J. Physiol. Endocrinol. Metab. 301, E901–911.
13
14
15 881 Peng, Y., Xiang, H., Chen, C., Zheng, R., Chai, J., Peng, J. & Jiang, S. (2013) MiR-224 impairs
16
17 882 adipocyte early differentiation and regulates fatty acid metabolism. The International
18
19 883 Journal of Biochemistry & Cell Biology 45, 1585–1593.
20
21
22 884 Quinlan, A.R. & Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic
23
24 885 features. Bioinformatics 26, 841–842.
25
26
27 886 Roberts, A., Pimentel, H., Trapnell, C. & Pachter, L. (2011a) Identification of novel transcripts
28
29 887 in annotated genomes using RNA-Seq. Bioinformatics 27, 2325–2329.
30
31 888 Roberts, A., Trapnell, C., Donaghey, J., Rinn, J.L. & Pachter, L. (2011b) Improving RNA-Seq
32
33 889 expression estimates by correcting for fragment bias. Genome Biol. 12, R22.
34
35
36 890 Ropka-Molik, K., Zukowski, K., Eckert, R., Gurgul, A., Piórkowska, K. & Oczkowicz, M. (2014)
37
38 891 Comprehensive analysis of the whole transcriptomes from two different pig breeds
39
40 892 using RNA-Seq method. Anim. Genet. 45, 674–684.
41
42
43 893 Rosen, E.D. & MacDougald, O.A. (2006) Adipocyte differentiation from the inside out. Nat.
44
45 894 Rev. Mol. Cell Biol. 7, 885–896.
46
47
48 895 Russell, T.D., Palmer, C.A., Orlicky, D.J., Bales, E.S., Chang, B.H.-J., Chan, L. & McManaman,
49
50 896 J.L. (2008) Mammary glands of adipophilin-null mice produce an amino-terminally
51
52 897 truncated form of adipophilin that mediates milk lipid droplet formation and secretion.
53
54
55 898 J. Lipid Res. 49, 206–216.
56
57
58
59
60

- 899 Russo, V., Buttazzoni, L., Baiocco, C., Davoli, M.R., Nanni Costa, N.L., Schivazappa, O.C. &
900 Virgili, P.C. (2000) Heritability of muscular cathepsin B activity in Italian Large White pigs.
901 Journal of Animal Breeding and Genetics 117, 37–42.
- 902 Russo, V., Fontanesi, L., Scotti, E., Beretti, F., Davoli, R., Nanni Costa, L., Virgili, R. &
903 Buttazzoni, L. (2008) Single nucleotide polymorphisms in several porcine cathepsin
904 genes are associated with growth, carcass, and production traits in Italian Large White
905 pigs. Journal of Animal Science 86, 3300–3314.
- 906 Serlachius, M. & Andersson, L.C. (2004) Upregulated expression of stanniocalcin-1 during
907 adipogenesis. Exp. Cell Res. 296, 256–264. Sethi, J.K. (2010) Activatin' human adipose
908 progenitors in obesity. Diabetes 59, 2354–2357.
- 909 Sethi, J.K. (2010) Activatin' Human Adipose Progenitors in Obesity. Diabetes 59, 2354–2357.
- 910 Smith, S.H., Wilson, A.D., Van Ettinger, I., MacIntyre, N., Archibald, A.L. & Ait-Ali, T. (2014)
911 Down-regulation of mechanisms involved in cell transport and maintenance of mucosal
912 integrity in pigs infected with *Lawsonia intracellularis*. Vet. Res. 45, 55.
- 913 Sodhi, S.S., Song, K.-D., Ghosh, M., Sharma, N., Lee, S.J., Kim, J.H., Kim, N., Mongre, R.K.,
914 Adhikari, P., Kim, J.Y., Hong, S.P., Oh, S.J. & Jeong, D.K. (2014) Comparative
915 transcriptomic analysis by RNA-seq to discern differential expression of genes in liver
916 and muscle tissues of adult Berkshire and Jeju Native Pig. Gene 546, 233–242.
- 917 Spurlock, M.E. & Gabler, N.K. (2008) The development of porcine models of obesity and the
918 metabolic syndrome. J. Nutr. 138, 397–402.
- 919 Toedebusch, R.G., Roberts, M.D., Wells, K.D., Company, J.M., Kanosky, K.M., Padilla, J.,
920 Jenkins, N.T., Perfield, J.W., Ibdah, J.A., Booth, F.W. & Rector, R.S. (2014) Unique
921 transcriptomic signature of omental adipose tissue in Ossabaw swine: a model of
922 childhood obesity. Physiol. Genomics 46, 362–375.

1
2
3 923 Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L.,
4
5 924 Rinn, J.L. & Pachter, L. (2012) Differential gene and transcript expression analysis of
6
7 925 RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7, 562–578.
8
9
10 926 Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg,
11
12 927 S.L., Wold, B.J. & Pachter, L. (2010) Transcript assembly and quantification by RNA-Seq
13
14 928 reveals unannotated transcripts and isoform switching during cell differentiation. Nat.
15
16 929 Biotechnol. 28, 511–515.
17
18
19 930 Trayhurn, P. (2005) Endocrine and signalling role of adipose tissue: new perspectives on fat.
20
21 931 Acta Physiol. Scand. 184, 285–293.
22
23
24 932 Uenishi, H., Eguchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H., Okumura, N., Hamasima,
25
26 933 N. & Awata, T. (2004) PEDE (Pig EST Data Explorer): construction of a database for ESTs
27
28 934 derived from porcine full-length cDNA libraries. Nucl. Acids Res. 32, D484–D488.
29
30
31 935 Uenishi, H., Eguchi-Ogawa, T., Shinkai, H., Okumura, N., Suzuki, K., Toki, D., Hamasima, N. &
32
33 936 Awata, T. (2007) PEDE (Pig EST Data Explorer) has been expanded into Pig Expression
34
35 937 Data Explorer, including 10 147 porcine full-length cDNA sequences. Nucl. Acids Res. 35,
36
37 938 D650–D653.
38
39
40 939 Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman,
41
42 940 F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric
43
44 941 averaging of multiple internal control genes. Genome Biol. 3, RESEARCH0034.
45
46
47 942 Varga, O., Harangi, M., Olsson, I. a. S. & Hansen, A.K. (2010) Contribution of animal models
48
49 943 to the understanding of the metabolic syndrome: a systematic overview. Obes Rev 11,
50
51 944 792–807.
52
53
54 945 Vicario, M., Skaper, S.D., Negro, A. (2014) The small heat shock protein HspB8: role in
55
56 946 nervous system physiology and pathology. CNS Neurol Disord Drug Targets 13, 885–895.
57
58
59
60

- 1
2
3 947 Wang, R.N., Green, J., Wang, Z., Deng, Y., Qiao, M., Peabody, M., Zhang, Q., Ye, J., Yan, Z.,
4
5 948 Denduluri, S., Idowu, O., Li, M., Shen, C., Hu, A., Haydon, R.C., Kang, R., Mok, J., Lee, M.J.,
6
7 949 Luu, H.L. & Shi, L.L. (2014) Bone Morphogenetic Protein (BMP) signaling in development
8
9 950 and human diseases. *Genes Dis* 1, 87–105.
10
11 951 Yamada, K., Santo-Yamada, Y. & Wada, K. (2002) Restraint stress impaired maternal
12
13 952 behavior in female mice lacking the neuromedin B receptor (NMB-R) gene. *Neurosci.*
14
15 953 *Lett.* 330, 163–166.
16
17 954 Zhou, C., Zhang, J., Ma, J., Jiang, A., Tang, G., Mai, M., Zhu, L., Bai, L., Li, M. & Li, X. (2013)
18
19 955 Gene expression profiling reveals distinct features of various porcine adipose tissues.
20
21 956 *Lipids Health Dis* 12, 75.
22
23 957 Zhou, Z.-Y., Li, A.-M., Adeola, A.C., Liu, Y.-H., Irwin, D.M., Xie, H.-B. & Zhang, Y.-P. (2014)
24
25 958 Genome-Wide Identification of Long Intergenic Noncoding RNA Genes and Their
26
27 959 Potential Association with Domestication in Pigs. *Genome Biol Evol* 6, 1387–1392.
28
29
30
31
32
33 960
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supporting information

Additional supporting information may be found in the online version of this article.

Supplementary Tables are included in the file:

TranscriptomeILW_SupplementaryTables.xlsx

Table S1 - Primers and PCR condition used for the validation.

EXT: primer pairs used for the amplification of a larger PCR product

INT: primer pairs used for the creation of the standard curve and for the qRT-PCR analysis

Table S2 - Number of reads for each sample

For each sample is indicated the total raw reads sequenced, total clean reads after the trimming and length filters and total reads mapped to the reference genome. Reported values refer to reads as they were single end (total clean paired reads are half the value in the table). Respective percentages are shown in the last three columns.

Table S3 - Types of transcripts expressed in backfat tissue, according to the considered genome annotations.

Transcripts, associated to eight Cufflinks class codes (see http://cufflinks.cbc.umd.edu/manual.html#class_codes), were classified into three major informative groups.

Table S4 – Intergenic transcript annotations.

1
2
3 980 **Table S5 - Transcript coding potential predicted by Coding Potential Calculator (CPC) for**
4
5 981 **intergenic transcripts.**

6
7
8 982 Reliable noncoding: CPC score <-1

9
10 983 Noncoding: CPC score $-1 \Rightarrow / \leq 0$

11
12 984 Coding: CPC score >0

13
14
15
16 985 **Table S6 - Most expressed transcripts (top 75%) detected in porcine backfat.**

17
18
19 986

20
21
22 987 **Table S7 - David functional annotation clustering of the most expressed genes.**

23
24 988 The 10 most relevant clusters are reported

25
26
27 989

28
29
30 990 **Table S8 - List of differentially expressed genes detected by Cuffdiff2.**

31
32
33 991

34
35
36 992 **Table S9 - List of differentially expressed genes detected by DeSeq2.**

37
38
39 993

40
41
42 994 **Table S10 - List of differentially expressed transcripts detected by Cuffdiff2.**

43
44
45 995

46
47
48 996 **Table S11 - Transcript coding potential predicted by Coding Potential Calculator (CPC) for**
49
50 997 **the differentially expressed transcripts.**

51
52
53 998 Reliable noncoding: CPC score <-1

54
55 999 Noncoding: CPC score $-1 \Rightarrow / \leq 0$

56
57
58 1000 Coding: CPC score >0

1
2
3 1001
4
5
6 1002 **Figure S1 - Read processing and alignment results.**
7
8
9
10 1003 **File: TranscriptomeLW_FigureS1.jpg**
11
12 1004 (A) The boxplots show the distribution of the reads considered in different steps and filters
13
14 1005 of the computational analysis pipeline, in the 20 considered samples. From left to right we
15
16 1006 show the number of raw reads sequenced, of clean reads resulted from the filtering steps,
17
18 1007 of reads successfully mapped to the reference genome, and of reads with unique alignment
19
20 1008 in the genome. (B) From the left, the bars show the average amounts, in the 20 considered
21
22 1009 samples, of reads spliced, aligned to an exon, to an intron, to intergenic regions (according
23
24 1010 to the *Sus scrofa* 10.2 genome annotation), or spanning exon-intron borders. Different
25
26 1011 colors indicate the proportion of read aligning to chromosomes (blue), genome scaffolds
27
28 1012 (red) or mitochondrial genome (yellow). (C) Number of expressed genes detected in
29
30 1013 different chromosomes, in mitochondrial genome (Mt) or in genome scaffolds (S).
31
32
33
34
35
36
37 1014 **Figure S2 – Gene expression distribution in FAT and LEAN groups.**
38
39
40 1015 **File: TranscriptomeLW_FigureS2.jpg**
41
42 1016 Cumulative gene expression is shown for the two groups. The figure represents the number
43
44 1017 of genes (horizontal axis) required to reach different percentages (vertical axis) of the
45
46 1018 overall gene expression. The inner panel focus on the cumulative expression curves for 50%
47
48 1019 and 75% of the expression.
49
50
51
52
53
54
55
56
57
58
59
60

1020 **Figure S3 – Alignment of the four detected isoforms of *PLIN2* gene (red box) with the**
1021 **porcine and vertebrates transcripts present in Ensembl.**

1022 **File: TranscriptomeLW_FigureS3.jpg**

1023

1024 **Figure S4 – Principal component analysis (PCA) based on gene expression profiles.**

1025 **File: TranscriptomeLW_FigureS4.jpg**

1026 The figure presents sample separation according to the two principal components,
1027 explaining most of the gene expression variation in the data. Samples are represented by
1028 dots, with green and orange colours indicating LEAN and FAT samples, respectively in Panel
1029 A) and red and blue indicating females and castrated males in Panel B). The PCA shows a
1030 clear separation of LEAN and FAT samples, with no separation of samples by sex.

1031

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure captions

Figure 1 – Transcripts and isoforms classification.

TranscriptomeLLW_Figure1.jpg

(A) Expressed transcript were classified, according to current gene annotations, into 8 types, reported with different colors (see legend) and grouped in three categories: K (known) collects transcripts found in reference annotation (yellow); I (isoform) collects alternative forms of transcripts (red shades); N collects new transcripts from not-annotated loci (green shades). The pie chart shows the number of transcripts detected, for each type, and their mutual proportions. Three transcript types of the N group have few elements (43 intronic; 5 possible polymerase run-on fragments; 3 transcript intron overlap a reference intron on the opposite strand) and are barely visible in the chart. (B) Transcript length distributions in the three categories. (C) Transcript expression level distribution for the three categories. (D) Number of genes (vertical axis) with their number of transcript isoforms detected (horizontal axis). Genes with only one transcript isoforms detected are the most frequent; however, genes with up to 31 different isoforms were detected. (E) The proportion of each transcript type for the transcript isoforms grouped as in (D). Genes with only one isoform (first bar) are mainly intergenic genes (green part). For genes having more than one isoform expressed, the proportion of novel isoforms detected increases along with the number of different isoforms for a gene (red part).

1052 **Figure 2 - Coding potential of new intergenic transcripts.**

1053 **TranscriptomeILW_Figure2.jpg**

1054 According to CPC scores, calculated both for the forward and for the reverse complement
1055 sequence, the intergenic transcripts were classified as “coding”, “non-coding” and “reliable
1056 non-coding”. (A) The pie chart shows numbers and proportions of intergenic transcripts
1057 falling in each category and provides the color code for the figure panels. (B) and (C) show
1058 respectively the distribution of lengths and of expression levels of intergenic transcripts,
1059 binned in the three categories. (D) Percentages of transcripts per category are compared,
1060 considering all the intergenic transcripts and the subset of the intergenic transcripts ranked
1061 within the 10% most expressed transcripts considering the whole transcriptome.

1062

1063 **Figure 3 – Differentially expressed genes and transcripts identified.**

1064 **TranscriptomeILW_Figure3.jpg**

1065 (A) Intersection of genes resulting differentially expressed (DE) according to DESeq2 and
1066 Cuffdiff2 analysis, and genes with at least one transcript resulting DE according to the
1067 transcript-level Cuffdiff2 analysis. We focused on the transcripts belonging to the 85 genes
1068 commonly identified by all the methods. (B) Proportions of the new and known DETs
1069 resulting higher- and lower-expressed in FAT vs. LEAN samples. (C) Number of DE genes
1070 mapping to chromosomes or to genome scaffolds (S).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1071

1072

1073

1074

1075

1076

1077

1078

1079

1080

1081

1082

1083

1084

1085

1086

1087

1088

1089

1090

1091

1092

Figure 4 - qRT-PCR validation of eleven genes differentially expressed according to RNA-seq data.

TranscriptomeLW_Figure4.jpg

(A) Log₂ FC values obtained from RNA-seq, according to Cuffdiff2 estimates, (black bars) and from qRT-PCR data (grey bars), for the eleven tested genes; (B) scatterplot showing the good correlation between the Log₂ FC values calculated with the two experimental methods.

Figure 5 – Genes differentially expressed between FAT and LEAN animals impact on specific and connected biological processes.

TranscriptomeLW_Figure5.jpg

Genes differentially expressed in FAT vs. LEAN pigs converge to specific functions that are more activated or impaired in FAT pigs. Genes and functions upregulated and downregulated in FAT pigs are shown in red and green shades, respectively. Several genes more expressed in FAT pigs are linked to fat deposition and lipid metabolism, to adipocyte differentiation and maturation or to signaling pathways regulating them; FAT pigs show as well increased expression of genes involved in inflammation and immunity and increased expression of genes involved in the control of complex behavior, also by inflammation-mediated secretory activity of adipocytes. Metabolic alterations induce chronic stress in the adipose tissue. FAT pigs shows under-expression of several genes involved in stress response by unfolded protein binding and misfolded protein aggregation prevention. The

1
2
3 1093 impairment of these functions might in turn augment inflammation and the consequent
4
5 1094 secretory activity and possibly induce senescence.
6
7
8 1095
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Tables

Table 1 - Genetic indexes and phenotypes for BFT and hot carcass weight of the pigs selected for the transcriptome analysis.

Group	Sample ID	Sex	Day of slaughter	Slaughter weight (kg) (*)	BFT phenotype (mm)	BFT EBV		
							Mean	SD
FAT	477	M	6	120	43	7.36	5.22	1.3
	476	F	6	119	37	7.17		
	474	M	2	113	38	6.03		
	482	F	9		42	5.75		
	478	F	7	118	33	5.05		
	516	F	3	115	36	4.88		
	479	M	8		41	4.76		
	483	F	10	119	38	4.41		
	489	M	18	108	35	3.54		
	484	M	15	128	35	3.27		
LEAN	490	M	19	113	24	-6.46	-8.63	1.4
	473	F	2	132	23	-7.54		
	487	M	18	110	23	-7.61		
	517	M	4	117	20	-7.71		
	485	F	17	126	20	-7.82		
	475	M	5	119	20	-8.03		
	481	M	9		22	-9.91		
	486	F	17	123	19	-10.27		
	488	F	18	128	19	-10.37		
	480	F	9		16	-10.59		

EBV: estimated breeding value

BFT: backfat thickness.

(*) slaughter weight: the hot carcass slaughter weight is reported. For four animals the weight is not available due to a problem of the automatic recording system at the slaughterhouse.

Table 2 - List of the DE genes and transcripts.

Cufflinks transcript ID	Cufflinks gene ID	Gene locus	Gene symbol	Cuffdiff2 gene log2 FC FAT vs. LEAN	Transcript group	Coding potential
TCONS_00102010	XLOC_040987	JH118612.1:113132-140205	<i>DSC2</i>	2.55	Known	-
TCONS_00061823	XLOC_023331	4:78928264-78930654	-	2.46	New	NON CODING
TCONS_00033774	XLOC_013001	15:140797584-140847461	<i>NYAP2</i>	2.38	New	CODING
TCONS_00061359	XLOC_023211	4:35670339-35685878	<i>DCSTAMP</i>	2.23	Novel isoform	CODING
TCONS_00095554	XLOC_036823	GL893451.1:11131-27485	<i>CRLF2</i>	2.21	Known	-
TCONS_00093244	XLOC_035190	9:50996895-51001264	-	2.17	New	NON CODING
TCONS_00087029	XLOC_032796	8:140307937-140315415	<i>SPP1</i>	2.09	Known	-
TCONS_00003007	XLOC_000806	1:283547172-283552108	-	2.07	New	CODING
TCONS_00095549	XLOC_036822	GL893451.1:7060-10625	-	2.03	New	NON CODING
TCONS_00067029	XLOC_025404	5:36179189-36186325	<i>LYZ</i>	2.03	Known	-
TCONS_00042581	XLOC_016514	18:6731368-6733669	<i>GIMAP2</i>	1.98	Known	-
TCONS_00061600	XLOC_023265	4:55660234-55715444	<i>ATP6V0D2</i>	1.96	Novel isoform	CODING
TCONS_00039556	XLOC_015432	17:53815353-53827092	<i>MMP9</i>	1.92	Known	-
TCONS_00039900	XLOC_015518	17:4110395-4192029	<i>MSR1</i>	1.92	Known	-
TCONS_00061643	XLOC_023283	4:62172539-62226917	<i>STMN2</i>	1.85	Known	-
TCONS_00034645	XLOC_013236	15:62409564-62414328	-	1.84	New	RELIABLE NON CODING
TCONS_00091509	XLOC_034399	9:63158999-63198155	<i>ST14</i>	1.79	Novel isoform	CODING
TCONS_00098750	XLOC_038994	GL895411.1:0-1073	<i>INHBB</i>	1.65	New	CODING
TCONS_00022322	XLOC_008474	13:32323641-32330286	<i>CCR1</i>	1.63	Known	-
TCONS_00044383	XLOC_017319	2:11807281-11850646	<i>MPEG1</i>	1.63	Known	-
TCONS_00075056	XLOC_028007	6:70039585-70099223	<i>PADI2</i>	1.6	Known	-

TCONS_00095875	XLOC_037025	GL893645.1:0-307	-	1.57	New	RELIABLE NON CODING
TCONS_00084869	XLOC_032187	8:71288921-71302169	AMBN	1.56	Known	-
TCONS_00033691	XLOC_012975	15:133452328-133456736	SLC11A1	1.56	Known	-
TCONS_00089513	XLOC_033895	9:90266412-90348498	SCIN	1.55	Known	-
TCONS_00042660	XLOC_016535	18:8306789-8313120	-	1.52	New	CODING
TCONS_00059834	XLOC_022860	4:99905518-99915176	CD1A	1.52	Novel isoform	CODING
TCONS_00059837	XLOC_022860	4:99905518-99915176	CD1A	1.52	Known	-
TCONS_00093519	XLOC_035465	9:101443296-101443885	GNPMB	1.46	New	NON CODING
TCONS_00098157	XLOC_038614	GL894967.1:126-517	GNPMB	1.42	New	CODING
TCONS_00018804	XLOC_007247	12:23439824-23441829	-	1.4	New	CODING
TCONS_00103084	XLOC_041497	X:37303173-37393818	CYBB	1.38	Known	-
TCONS_00065337	XLOC_024931	5:52504178-52625145	BCAT1	1.37	Novel isoform	CODING
TCONS_00098113	XLOC_038589	GL894923.1:47-563	GNPMB	1.36	New	CODING
TCONS_00002441	XLOC_000664	1:227333991-227356844	PLIN2	1.32	Novel isoform	CODING
TCONS_00044392	XLOC_017322	2:12191483-12243400	LPXN	1.31	Known	-
TCONS_00084565	XLOC_032101	8:33970571-33982450	UCHL1	1.27	Novel isoform	CODING
TCONS_00067389	XLOC_025495	5:64579162-64590512	OLR1	1.26	Known	-
TCONS_00059747	XLOC_022835	4:97720982-97736619	CD48	1.25	Known	-
TCONS_00028769	XLOC_011055	14:143745489-143752509	GMFG	1.23	Known	-
TCONS_00029056	XLOC_011139	14:8804077-8816800	STC1	1.23	Novel isoform	CODING
TCONS_00098643	XLOC_038938	GL895339.1:13269-61205	COTL1	1.15	Known	-
TCONS_00100592	XLOC_040068	GL896326.1:1999-3913	ACP5	1.13	Known	-
TCONS_00096837	XLOC_037668	GL894123.1:0-400	CD163	1.13	New	CODING
TCONS_00097297	XLOC_037990	GL894401.1:0-471	CD163	1.13	New	CODING
TCONS_00005002	XLOC_001331	1:125897935-125953413	AQP9	1.09	Known	-
TCONS_00096863	XLOC_037686	GL894145.1:0-401	CD163	1.09	New	CODING
TCONS_00071337	XLOC_027094	6:74616232-74621248	C1QC	1.08	Known	-

TCONS_00012469	XLOC_005058	11:21534980-21685851	<i>LCP1</i>	1.07	Novel isoform	CODING
TCONS_00079920	XLOC_030238	7:94900207-94906867	<i>AKAP5,</i> <i>LOC100153460</i>	1.06	Novel isoform	CODING
TCONS_00041537	XLOC_016257	18:6613761-6621027	<i>GIMAP4</i>	1.06	Known	-
TCONS_00097908	XLOC_038444	GL894747.1:3047-10617	<i>HMOX1</i>	1.06	Novel isoform	CODING
TCONS_00030401	XLOC_011444	14:71516962-71521335	<i>EGR2</i>	1.05	Known	-
TCONS_00030878	XLOC_011579	14:117265093-117349965	<i>BLNK</i>	1.04	Known	-
TCONS_00056578	XLOC_021190	3:77408776-77439119	<i>PLEK</i>	1.04	Known	-
TCONS_00071335	XLOC_027093	6:74609911-74612993	<i>C1QA</i>	1.02	Known	-
TCONS_00081915	XLOC_030757	7:54395230-54406136	<i>BCL2A1</i>	1.01	Known	-
TCONS_00041554	XLOC_016261	18:6872940-6875292	<i>GIMAP1</i>	1	Known	-
TCONS_00085005	XLOC_032236	8:79743274-79751980	<i>SFRP2</i>	0.99	Known	-
TCONS_00098919	XLOC_039115	GL895590.1:0-1327	<i>GPNMB</i>	0.91	New	NON CODING
TCONS_00068526	XLOC_026077	5:52625315-52630242	<i>BCAT1</i>	0.89	New	RELIABLE NON CODING
TCONS_00062055	XLOC_023401	4:97099149-97103132	<i>FCER1G</i>	0.87	Known	-
TCONS_00009719	XLOC_003695	10:48841010-48961015	<i>MRC1</i>	0.86	Novel isoform	CODING
TCONS_00030894	XLOC_011584	14:117670639-117938624	<i>PIK3AP1</i>	0.85	Known	-
TCONS_00017526	XLOC_006800	12:36561025-36604089	<i>CLTC</i>	0.8	Novel isoform	CODING
TCONS_00062959	XLOC_023614	4:119674090-119703427	<i>CD53</i>	0.78	Known	-
TCONS_00081898	XLOC_030753	7:53623061-53644262	<i>CTSH</i>	0.78	Known	-
TCONS_00060570	XLOC_023035	4:119013307-119039899	<i>ADORA3</i>	0.74	Known	-
TCONS_00052401	XLOC_020144	3:11035819-11055510	<i>LAT2</i>	0.71	Known	-
TCONS_00004118	XLOC_001095	1:35133812-35137388	<i>CTGF</i>	0.68	Known	-
TCONS_00045043	XLOC_017499	2:59214054-59218018	<i>IFI30</i>	0.65	Known	-
TCONS_00004124	XLOC_001096	1:35240242-35281384	<i>ENPP1</i>	0.62	Known	-
TCONS_00062884	XLOC_023592	4:116704501-116707235	<i>OLFML3</i>	-0.54	Known	-
TCONS_00035484	XLOC_013426	15:131680309-131684630	<i>IGFBP5</i>	-0.65	Known	-

TCONS_00101718	XLOC_040809	JH118426.1:306724-312138	-	-0.77	New	RELIABLE NON CODING
TCONS_00063805	XLOC_024145	4:77261119-77264781	-	-0.77	New	NON CODING
TCONS_00050164	XLOC_018733	2:124815021-124828122	<i>CDO1</i>	-0.9	Novel isoform	CODING
TCONS_00101559	XLOC_040715	GL896532.1:212-2567	<i>ADSSL1</i>	-1.02	New	NON CODING
TCONS_00079927	XLOC_030240	7:94987617-94990126	<i>HSPA2</i>	-1.1	Known	-
TCONS_00083805	XLOC_031620	7:66542203-66555641	-	-1.18	New	CODING
TCONS_00041725	XLOC_016313	18:15292592-15295178	-	-1.61	New	CODING
TCONS_00048853	XLOC_018425	2:65175406-65180520	<i>DNAJB1</i>	-1.66	Novel isoform	CODING
TCONS_00029533	XLOC_011248	14:35688332-35701411	<i>HSPB8</i>	-1.81	Known	-
TCONS_00094194	XLOC_036009	GL892492.1:0-3540	<i>HSPA1B</i>	-2.32	New	NON CODING
TCONS_00101505	XLOC_040677	GL896522.1:9039-10877	<i>HSPA1A</i>	-2.57	New	RELIABLE NON CODING
TCONS_00098059	XLOC_038555	GL894890.1:5-696	<i>HSP70</i>	-3.44	New	NON CODING

Table 3 - David functional annotation clustering obtained considering the significant Biological Processes GO terms (Benjamini adjusted P-values <0.05) of genes more expressed in FAT than in LEAN animals.

Annotation Cluster 1		Enrichment Score: 7.0	
Term	Count	Genes	
GO:0006954~inflammatory response	12	<i>C1QA, SLC11A1, CYBB, ADORA3, OLR1, HMOX1, CCR1, LYZ, C1QC, BLNK, CD163, SPP1</i>	
GO:0006952~defense response	15	<i>ADORA3, OLR1, CCR1, LYZ, COTL1, C1QC, CD163, INHBB, CD48, C1QA, SLC11A1, CYBB, HMOX1, SPP1, BLNK</i>	
GO:0009611~response to wounding	14	<i>ADORA3, PLEK, OLR1, CCR1, LYZ, C1QC, CD163, C1QA, SLC11A1, CYBB, CTGF, HMOX1, SPP1, BLNK</i>	
GO:0009605~response to external stimulus	17	<i>ADORA3, PLEK, OLR1, CCR1, LYZ, C1QC, CD163, INHBB, C1QA, SLC11A1, CYBB, CTGF, SFRP2, HMOX1, STC1, SPP1, BLNK</i>	
GO:0050896~response to stimulus	29	<i>ADORA3, AQP9, ENPP1, CCR1, UCHL1, ACP5, C1QC, CD48, SLC11A1, PLIN2, CTGF, HMOX1, FCER1G, BLNK, SPP1, EGR2, OLR1, PLEK, LYZ, CD1A, COTL1, CD163, INHBB, C1QA, CYBB, LAT2, SFRP2, STC1, LCP1</i>	
GO:0006950~response to stress	19	<i>ADORA3, AQP9, PLEK, OLR1, CCR1, UCHL1, LYZ, COTL1, C1QC, CD163, INHBB, CD48, C1QA, SLC11A1, CYBB, CTGF, HMOX1, SPP1, BLNK</i>	
Annotation Cluster 2		Enrichment Score: 2.7	
Term	Count	Genes	
GO:0001775~cell activation	7	<i>CD48, SLC11A1, LAT2, PLEK, LCP1, BLNK, GIMAP1</i>	
GO:0002274~myeloid leukocyte activation	4	<i>CD48, SLC11A1, LAT2, GIMAP1</i>	
GO:0046649~lymphocyte activation	6	<i>CD48, SLC11A1, LAT2, LCP1, BLNK, GIMAP1</i>	
Annotation Cluster 3		Enrichment Score: 2.4	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Term	Count	Genes
GO:0048583~regulation of response to stimulus	10	<i>C1QA, SLC11A1, LAT2, PLEK, ENPP1, HMOX1, FCER1G, C1QC, SPP1, GIMAP1</i>
GO:0050776~regulation of immune response	7	<i>C1QA, SLC11A1, LAT2, HMOX1, FCER1G, C1QC, GIMAP1</i>
GO:0050778~positive regulation of immune response	6	<i>C1QA, SLC11A1, LAT2, FCER1G, C1QC, GIMAP1</i>
GO:0002443~leukocyte mediated immunity	5	<i>C1QA, SLC11A1, LAT2, FCER1G, C1QC</i>
GO:0002682~regulation of immune system process	8	<i>C1QA, SLC11A1, LAT2, HMOX1, SCIN, FCER1G, C1QC, GIMAP1</i>
Annotation Cluster 4	Enrichment Score: 2.0	
Term	Count	Genes
GO:0060348~bone development	6	<i>AMBN, CTGF, ACP5, STC1, GPNMB, SPP1</i>
GO:0031214~biomineral formation	4	<i>AMBN, ENPP1, GPNMB, SPP1</i>
GO:0001503~ossification	5	<i>AMBN, CTGF, STC1, GPNMB, SPP1</i>
GO:0001501~skeletal system development	7	<i>AMBN, CTGF, MMP9, ACP5, STC1, GPNMB, SPP1</i>
Annotation Cluster 5	Enrichment Score: 1.6	
Term	Count	Genes
GO:0001775~cell activation	7	<i>CD48, SLC11A1, LAT2, PLEK, LCP1, BLNK, GIMAP1</i>

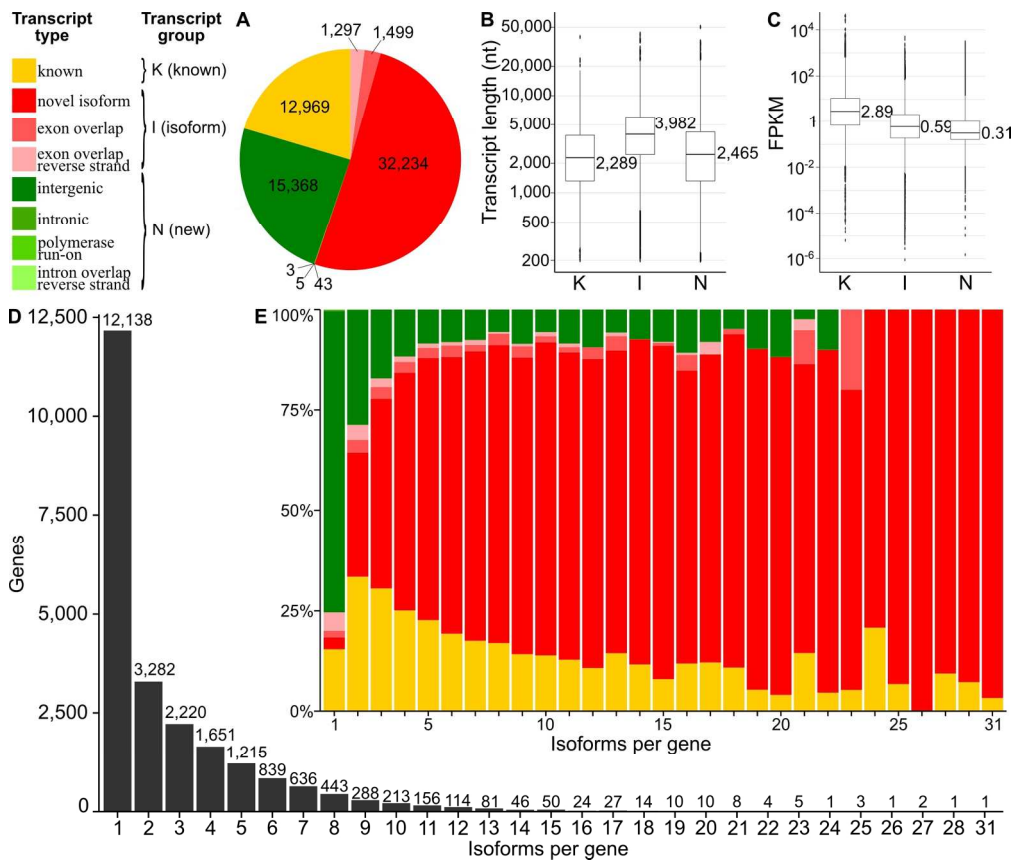


Figure 1

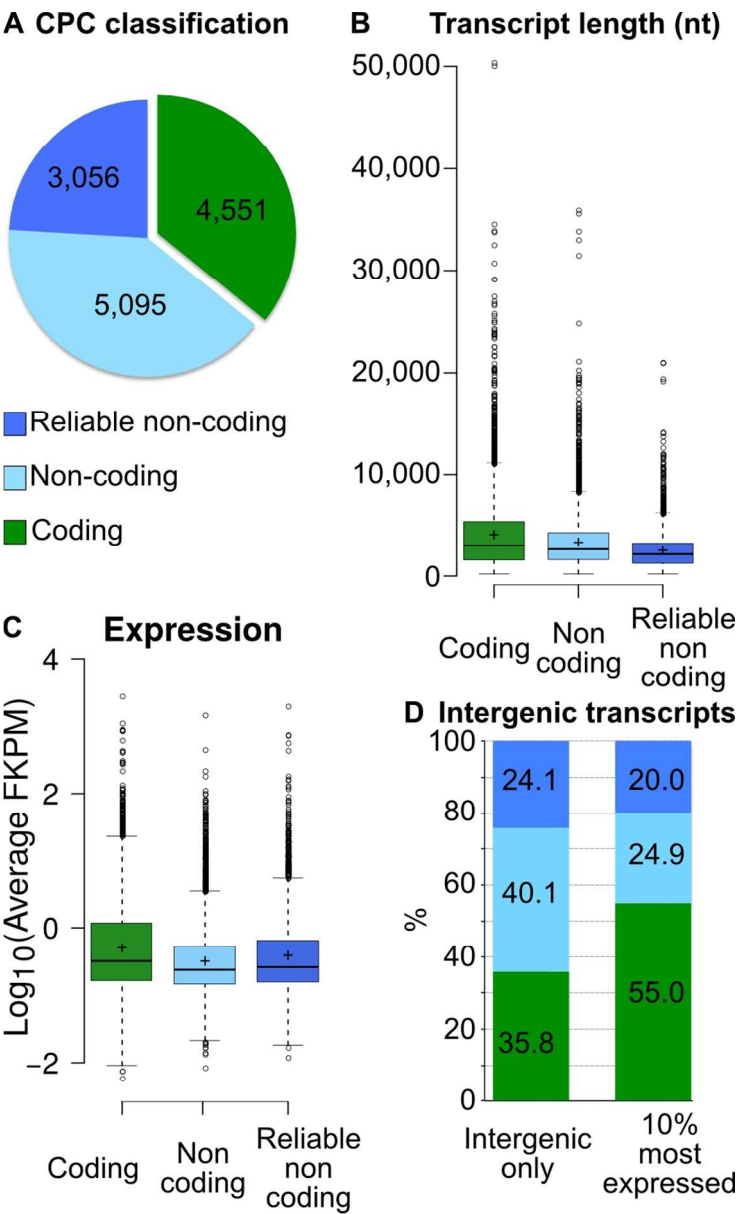


Figure 2

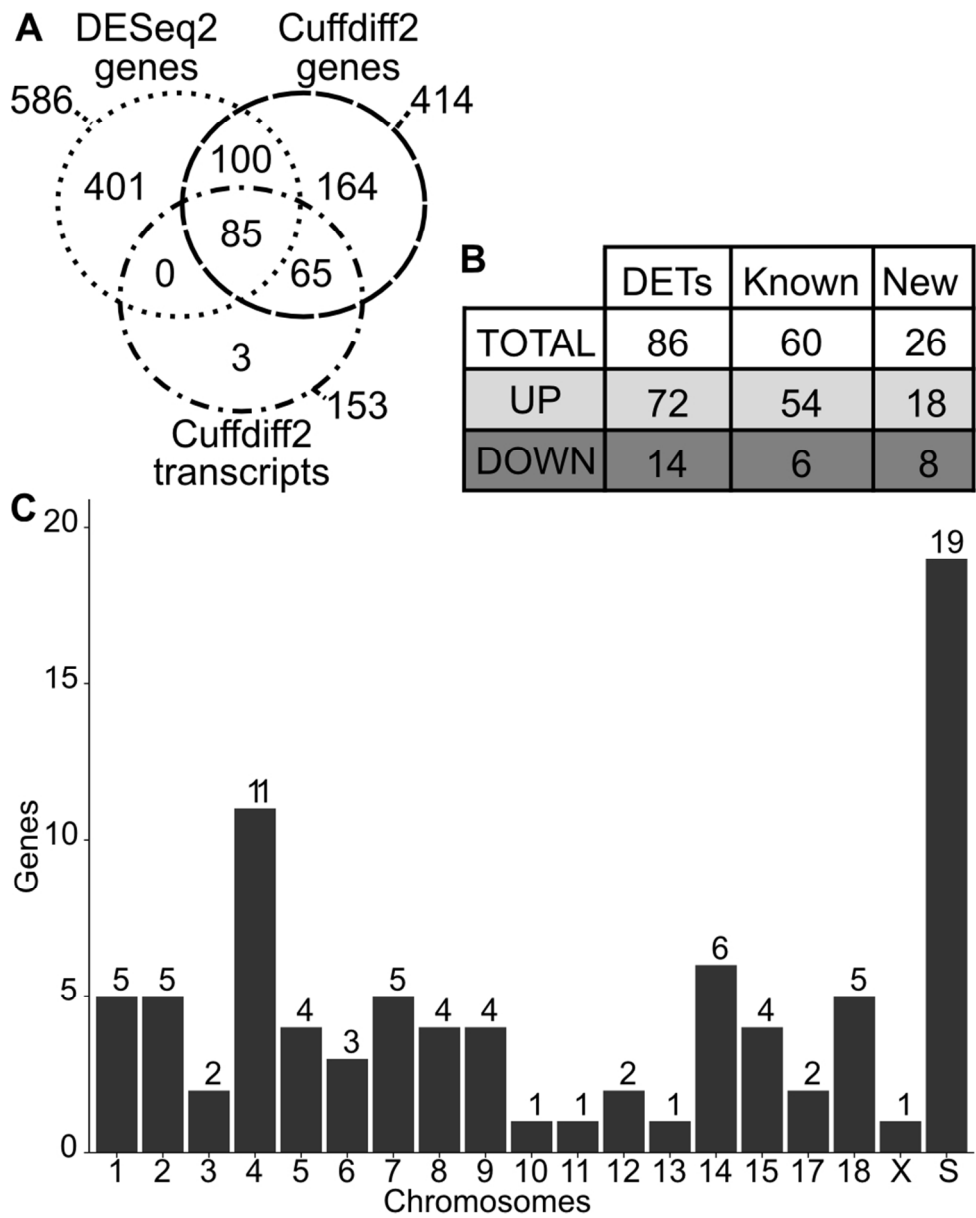


Figure 3
85x105mm (300 x 300 DPI)

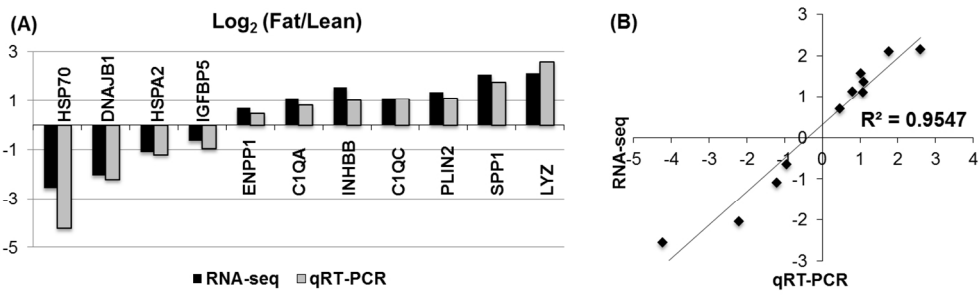


Figure 4
242x73mm (150 x 150 DPI)

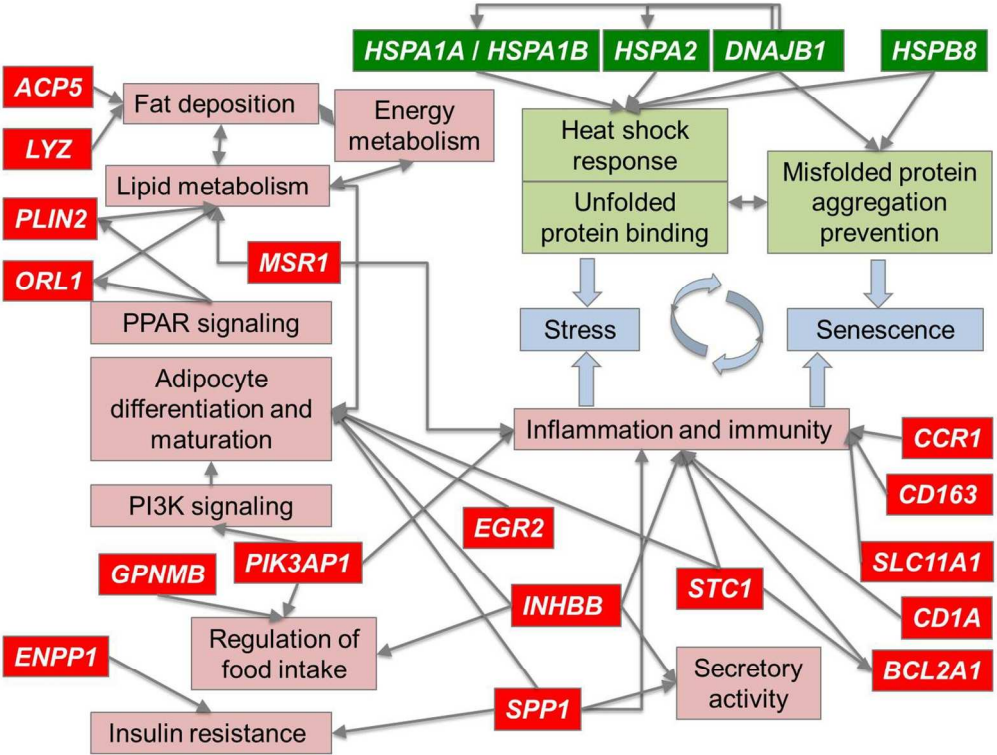


Figure 5
249x190mm (150 x 150 DPI)