1 Identification of Novel Genes and Variations Associated to

2 Glycolytic Potential Based on Pig Model

3 Wangjun Wu^{*,1,2}, Zengkai Zhang^{*,1}, Zhe Chao⁺, Bojiang Li^{*}, Caibo Ning^{*}, Aiwen

4 Jiang^{*}, Chao Dong^{*}, Wei Wei^{*}, Jie Chen^{*}, Honglin Liu^{*}

5 ^{*} Department of Animal Genetics, Breeding and Reproduction, College of Animal

6 Science and Technology, Nanjing Agricultural University, Nanjing, 210095, China

7 ⁺ Institute of Animal Sciences & Veterinary, Hainan Academy of Agricultural

8 Sciences, Haikou 571100, China

9 wangjunwu4062@163.com

10 2017105031@njau.edu.cn

11 chaozhe.cn@126.com

12 2015205004@njau.edu.cn

13 2016205009@njau.edu.cn

14 2015105077@njau.edu.cn

15 2016105025@njau.edu.cn

23

24 **Running title:**

25 Genes Affecting Glycolytic Potential

26 **Keywords:**

27 pig, transcriptome, glycolytic potential, glycogen, meat quality

28 **Corresponding author:**

29 Wangjun Wu

30 Department of Animal Genetics, Breeding and Reproduction, College of Animal

31 Science and Technology, Nanjing Agricultural University, Nanjing, 210095, China

32 **E-mail:**

33 wangjunwu4062@163.com; wuwangjun2012@njau.edu.cn

34

35 ABSTRACT

36 In livestock, glycolytic potential (GP) is a critical indicator for evaluating the meat
37 quality. To date, two major genes *protein kinase AMP-activated γ 3 non-catalytic*
38 *subunit gene (PRKAG3)* and *phosphorylase kinase catalytic subunit gamma*
39 *1(PHKG1)*, and corresponding cause mutations influencing GP have been confirmed
40 in pigs. Therefore, the aim of this study to identify the novel candidate genes and
41 variations related to GP-related traits using a four-hybrid pig model [Pietrain (P)×
42 Duroc (D)] ×[(Landrace) ×(Yorkshire)]. We totally constructed six RNA-seq libraries
43 using *longissimus dorsi* (LD) muscles, and each library contained two higher GP (H)
44 or two lower GP (L) individuals. A total of 525, 698 and 135 differentially expressed
45 genes (DEGs) were identified between H11 vs L11, H9 vs L9, and H5 vs L5 groups
46 using PossionDis method, respectively. Notably, we found 97 non-redundant DEGs
47 were mapped to GP related QTLs from three paired comparison groups. Moreover, 69
48 DEGs were identified between H (H11, H9 and H5) and L (L11, L9 and L5) groups
49 using NOIseq method. Additionally, 1,076 potential specific SNPs were figured out
50 between H and L groups, and approximately 40 large Indels with a length $\geq$ 5bp were

51 identified in each sequencing library. In conclusion, our data provide foundation for
 52 further confirming the key genes and the functional mutations affecting GP-related
 53 traits in pigs, and also pave the way for elucidating the underling molecular regulatory
 54 mechanisms of glycogen metabolism in future study. Moreover, this study might
 55 provide valuable information for study on human glycogen storage diseases.

56

57 INTRODUCTION

58 Pork is the most commonly consumed meat worldwide, its quality have received
 59 much attention due to the increase of inferior meat caused by high-strength selection
 60 for leaner pork products. However, improving meat quality is one of the most
 61 challenging tasks in pig breeding because that it hard to get effective genetic progress
 62 using traditional conventional breeding methods. Therefore, identification and
 63 application of major genes and variations have become the research hotspots in pig
 64 molecular breeding.

65 Meat is the product of skeletal muscle which is subjected to a serial of structural
 66 and biochemical changes after animal slaughter. During the transformation of muscle
 67 to meat, the speed and extent of pH value change postmortem are key factors
 68 affecting meat quality, such as water-holding capacity (WHC), tenderness, color *etc*
 69 (Bee et al. 2007; Popp et al. 2015; Werner et al. 2010). Notably, pH value is directly
 70 influenced by glycolysis, with glycogen being converted to lactate and H^+ . Therefore,
 71 glycogen, the substrate of glycolysis, its content largely determines the pH value
 72 change and closely related to many meat quality traits (Hamilton et al. 2003; Zybert et

73 al. 2013).

74 The quality of pork is determined by multiple of factors, including genetics,
 75 nutrition, feeding methods, slaughtering procedures, and handling of carcass(Jasinska
 76 and Kurek 2017; Lebret 2008; Pugliese and Sirtori 2012). Undoubtedly, genetics is
 77 the most essential factor affecting meat quality. Currently, *ryanodine receptor 1*
 78 (*RYR1*), *protein kinase AMP-activated γ 3 non-catalytic subunit gene* (*PRKAG3*; also
 79 known as *rendement napole* (*RN*)) are the major genes affecting meat quality. *RYR1*,
 80 is a calcium-dependent calcium release channel in the sarcoplasmic reticulum of
 81 skeletal muscle, and the g.1843C>T variation in the coding region has been confirmed
 82 as the cause mutation for Pale, Soft, and Exudative (PSE) meat (Fujii et al. 1991).
 83 *PRKAG3*, the first major gene affecting glycogen content, its R200Q mutation
 84 significantly increases $\sim 70\%$ muscle glycogen content in *RN* animals, which
 85 eventually lead to acid meat (Milan et al. 2000). Recently, the second glycogen
 86 related major gene, *phosphorylase kinase catalytic subunit gamma 1*(*PHKG1*), its
 87 mutation (g.8283C>A) in its ninth intron causes a 32 bp deletion in the open reading
 88 frame, generating a premature stop codon, which has been demonstrated to alter GP

(43%) and decreases WHC (>20%)(Ma et al. 2014). Notably, the effects of the cause mutations in *PRKAG3* and *PHKG1* were observed specifically in Hampshire and Duroc breeds, respectively. In addition, although many different QTLs related to glycogen and related traits have been deposited in pigQTLdb (<http://www.animalgenome.org/cgi-bin/QTLdb/SS/index>) and some candidate genes and variations related to GP have been investigated (Kaminski et al. 2010; Sieczkowska et al. 2010), the knowledge of major genes and variations related to glycogen and related traits are still limited.

Therefore, the objective of this study was to identify the potential candidate genes related to glycogen metabolism traits using RNA-seq technology with extremely GP phenotypic pigs, and find potential functional variations affecting GP phenotypes through SNP analysis. Our results could provide valuable references for identification of major genes and variations related to glycogen metabolism traits.

MATERIALS AND METHODS

Animals and ethics statement

A total of twelve experimental pigs derived from a 279 four-hybrid pigs which crossed by six F₁ (Pietrain × Duroc) boars and twenty five F₁ (Landrace× Yorkshire) sows were used for transcriptome analysis. The (P×D)×(L×Y) commercial pig population were reared indoors with uniform environment, and fed with a standard commercial feed ration with free access to water. All the pigs were slaughtered humanely in a standardized commercial abattoir (Jiangsu Sushi Group, China) at an average age of 176 days with an average body weight 110kg in eleven batches (No.1-11) from 1st to 2rd September, 2015. The number of pigs in each batch varied from 23 to 31 pigs. All procedures involving animals were approved by the Institutional Animal Care and Use Committee of Nanjing Agricultural University, Nanjing, Jiangsu, P.R. China (SYXK2011-0036).

Samples collection and glycolytic potential determination

The LD muscles from all the experimental individuals were excised, and snap frozen in liquid nitrogen at approximately 45min postmortem. Subsequently, the glycolytic potential was determined with 2g muscle tissue between 1th and 2th thoracic vertebrae from posterior. Muscle glucose (MG), glycogen (G), and lactate

120 content were determined using the glucose, glycogen and lactate test kit (Jiancheng,
 121 Nanjing), respectively. The glucose-6-phosphate (G6P) was determined using the
 122 Glucose-6-Phosphate Assay Kit (Sigma, Spruce Street, St. Louis, MO, USA).
 123 Residual glycogen (RG) was the combined content of glycogen and glucose.
 124 Glycolytic potential (GP) was calculated according to the formula proposed by Monin
 125 and Sellier (Monin and Sellier 1985). $GP = 2 \times ([\text{glycogen}] + [\text{glucose}] +$
 126 $[\text{glucose-6-phosphate}]) + (\text{lactate})$. GP was represented as micromoles of lactate
 127 equivalent per gram of fresh tissue. In addition, the carcass weight (CW), drip loss
 128 measured at 24h postmortem (DL_{24h}), drip loss measured at 48h postmortem (DL_{48h}),
 129 cooking loss (CL), and pH values at 45min and 24h postmortem were measured
 130 according the international and national standard protocol. In addition, based on the
 131 results of GP phenotypic values, twelve experimental pigs from three slaughtered
 132 batches (No. 5, 9 and 11) with the most significant differences between H and L
 133 groups were selected for further experiment. A total of twelve economic traits were
 134 recorded, and the phenotypic significant differences were statistically analyzed in
 135 IBM SPSS 20.0 using unpaired Student's *t*-test method.

Total RNA extraction and library construction

The LD muscles from twelve experimental pigs derived from three slaughtered batches (No. 5, 9 and 11), four individuals for each batch (two with the higher GP and two with the lower GP) were selected for construction of transcriptome libraries, and the schematic overview of RNA-seq libraries construction was shown in **Figure S1**. Total RNA of each individual was extracted using the TRIzol reagent (Life Technologies, Carlsbad, CA), separately. The purity and quantity of total RNA were firstly detected using NanoDrop 2000 spectrophotometer (Thermo Scientific), while integrity was assessed using Agilent 2100 (Agilent, Technologies, Inc.). Only the RNA with OD 260/280 between 1.80-2.20 and RIN $\geq$ 8.0 were used for libraries construction. Briefly, the total RNA from two individuals with the higher GP and two with the lower GP at each slaughtered batch were pooled equally for a sequencing library, respectively. Then, the sequencing libraries were prepared with mRNA-Seq Sample Preparation Kit (Illumina, San Diego, CA, USA) according to the protocol described in the document, and the quality of libraries were validated through checking the size, purity, and the concentration on an Agilent Technologies 2100

Bioanalyzer (Agilent, Technologies, Inc.). Accordingly, a total of six libraries were constructed, and named H11 and L11, H9 and L9, and H5 and L5, separately. ‘H’ and ‘L’ represent the libraries were derived from the higher and the lower GP individuals, respectively.

Sequencing and raw data processing and alignment analysis

Sequencing was conducted on HiSeq 4000 platform, and 150 bp paired-end reads were produced. The raw data generated were suffered from a serial of filtration, including the adapter, invalid reads containing poly-N and low-quality reads. The remaining high quality reads after filtering were considered as clean reads and saved as FASTQ format. Furthermore, the clean reads were aligned to *Sus scrofa* reference genome (*Sus scrofa* 10.2) using HISAT v0.1.6 (Kim et al. 2015).

Bioinformatics analysis

Subsequently, transcripts were reconstructed using StringTie software v1.0.4 (Pertea et al. 2015) and novel transcripts were predicted with the genome annotation information using cuffcompare tool (Trapnell et al. 2012), and the differentially

splicing genes (DSG) between samples were detected using rMATS v3.0.9 (Shen et al. 2014) and False Discovery Rate (FDR) ≤ 0.05 was set as the threshold value to judge the significance of DSG. After that, the novel coding transcripts were integrated into known transcripts to obtain a complete reference sequence for gene expression analysis. Then, the clean reads were mapped to the reference sequence using Bowtie2 v2.2.5 (Langmead and Salzberg 2012) and expression levels of genes and transcripts were calculated using RSEM v1.2.12 (Li and Dewey 2011).

Furthermore, the differentially expressed genes (DEGs) between the paired samples from the same batch were detected using PoissonDis method based on the poisson distribution (Audic and Claverie 1997), and the corrected *P*-value, FDR ≤ 0.001 and absolute value of log₂ Foldchange ≥ 1 were considered as DEGs. While, the DEGs between the higher (H) and the lower (L) groups considering biological replicates were detected using NOIseq method based on noisy distribution (Tarazona et al. 2011), and the corrected *P*-value ≥ 0.7 and absolute value of log₂ Foldchange ≥ 1 were considered as DEGs.

Thereafter, the DEGs from PoissonDis method were mapped to QTL with the

pig QTL database (<http://www.animalgenome.org/cgi-bin/QTLdb/SS/index>) and the QTLs related to glycolytic potential were refined. Additionally, the single nucleotide polymorphisms (SNPs) and insertion-deletion variations (Indels) in each sequencing library were detected with genome mapping result using GATK software v3.4 (McKenna et al. 2010), and the potential differentially and specifically SNPs and Indels were refined manually through comparison analysis between ‘H’ and ‘L’ libraries using the summarized list data of total SNPs (**File S1**). Moreover, the differentially expressed genes containing the potential specifically SNPs were figured out according gene ID and chromosome position of SNPs using SNP annotation information list data (**File S1**) and DEGs annotation information list data (**File S2**).

Validation of DEGs and SNPs

To validate the reliability of sequencing data, real-time PCR was used to confirm the DEGs. Total RNA were extracted from twelve LD muscle samples for transcriptome sequencing using TRIzol reagent (Invitrogen, Life technologies, Carlsbad, USA), respectively, The first strand cDNA was synthesized using Prime

ScriptTM RT reagent (TaKaRa, Dalian, China), and real-time RT-PCR was conducted on a Step-One Plus Real-Time PCR System (Applied Biosystems, Carlsbad, CA, USA) using the AceQ qPCR SYBR Green Master Mix (Vazyme, Nanjing, China). A total of nine DEGs were selected for quantitative analysis and all the quantitative primers used in this study were shown in **Table S1**. All reactions were performed in four repetitions and the reference gene *HPRT* was used to normalize gene expression levels. Relative gene expression levels were calculated using the comparative Ct ($\Delta\Delta Ct$) value method. All statistical analyses were performed in IBM SPSS 20.0, and unpaired Student's *t*-test was used to evaluate the statistical significance of differences. Moreover, correlation analysis was conducted between the fold change values of DEGs from transcriptome sequencing and real-time PCR with Pearson correlation coefficient, $P < 0.05$ was considered statistically significant difference. Additionally, eight potential differentially SNPs were validated through PCR product sequencing method, and the information of PCR amplification primers and SNPs location were shown in **Table S2**.

213 **Data availability**

214 The raw data of transcriptome sequencing data have been deposited to NCBI Short
 215 Read Archive (<https://www.ncbi.nlm.nih.gov/sra>) and are accessible through
 216 accession number (SRP127540). **Table S1** contains quantitative primers used in this
 217 study. **Table S2** contains sequencing primers and SNPs location information in this
 218 study. **Table S3** provides phenotypic data for all experimental pigs with extreme GP
 219 value. **Figure S1** is the schematic overview of RNA-seq libraries construction. **Figure**
 220 **S2** depicts the distribution of base quality on clean reads. **File S1** contains the
 221 potential differentially and specifically SNPs information identified in this study. **File**
 222 **S2** provides DEGs from PossionDis method. **File S3** contains novel predicted
 223 coding_transcripts. **File S4** contains differentially splicing genes. **File S5** provides the
 224 genes expression data for all sample libraries. **File S6** provides DEGs from NOIseq
 225 method. **File S7** provides the glycolytic potential related QTLs location of DEGs from
 226 PossionDis method. **File S8** contains the Indel information identified in this study.
 227 **File S9** provides the information on more than 5bp Indel.

228 **RESULTS**

229 **Phenotypes**

230 In this study, twelve pigs from three batches (four individuals for each batch)
 231 represent the higher GP groups (H group: H11, H9, and H5) and the lower GP groups
 232 (L group: L11, L9, and L5). H11 and L11, H9 and L9, and H5 and L5 were derived
 233 from the same batch, respectively. The phenotypic characteristics of each
 234 experimental pig were shown in **Table S3**. From the results, the GP phenotypic values
 235 between H and L group individuals were obviously different (**Table 1**). GP phenotypic
 236 mean values between H5 and L5, and H9 and L9 were reached extremely significant
 237 levels ($P < 0.01$). Although the difference of GP phenotypic mean values between H11
 238 and L11 did not reach statistic significant level due to the large variation in H11 group
 239 individuals, the obvious difference was actually existed between H11 and L11. In
 240 addition, the significant differences of RG, MG, LA, pH_{45min}, DL_{48h} and CL between
 241 H and L group were also observed.

242 **Sequencing data and alternative splicing analysis**

243 In this study, six sequencing libraries (H11, H9, H5 and L11, L9, L5) were

constructed (**Figure S1**) and sequenced on an HiSeq 4000 platform. More than 6 G bases data was generated from each library, representing more than 50.0 million raw reads. The output of data and alignment results were summarized in **Table 2**. After filtering, about 45.0 million clean reads were obtained, the average ratio of clean reads reached 85.0%, and the ratio of clean reads Q20 (those with a base quality > 20 and error rate < 0.01) was higher than 97%, indicating that high-quality clean reads were obtained from each library (**Figure S2**). Moreover, the mapping analysis of clean reads showed that more than 29.0 million clean reads were mapped to reference genome for each library, the ratio reached to 65.0%. In the present study, about 14,000 known genes and hundreds of novel genes were identified in each library, representing approximate 20,000 known transcripts and 7,000 novel transcripts (**Table 2**). The sequence information of novel transcripts can be found in **File S3-1 and File S3-2**. In addition, alternative splicing (AS) analysis showed that five types of AS events were detected, including alternative 3' splice site (A3SS), alternative 5' splice site (A5SS), skipping exon (SE), mutually exclusive exon (MXE), and retained intron (RI). The differentially splicing genes (DSG) between samples were shown in

File S4.

Identification of DEGs

Expression levels of genes and transcripts described by Fragments Per Kilobase of transcript per Million fragments mapped (FPKM) in each sample were calculated, the results were shown in **File S5**. On this basis, the DEGs with significance between H11 vs L11, H9 vs L9, and H5 vs L5 were analyzed using PossionDis method. The number of DEGs identified between H11 vs L11, H9 vs L9, and H5 vs L5 were 525, 698 and 135 with $FDR \leq 0.001$ and $|\log_2 \text{Foldchange}| \geq 1$, and the volcanoplots and statistic of DEGs was shown in **Figure 1** and detailed information of DEGs was shown in **File S2**. Moreover, the overlapped DEGs among three groups were analyzed, and 27 DEGs were overlapped between H11 vs L11 (HL11) and H5 vs L5 (HL5), 26 DEGs were overlapped between H9 vs L9 (HL9) and H5 vs L5 (HL5), 157 DEGs were overlapped between H11 vs L11 (HL11) and H9 vs L9 (HL9), while only 6 DEGs were overlapped among three groups (**Figure 2**). In addition, the DEGs between H and L groups considering biological replicates were detected using

NOIseq method. A total of 69 DEGs with the corrected P -value ≥ 0.7 and $|\log_2$
Foldchange ≥ 1 were detected in this study (**File S6**). Moreover, 36 DEGs from
NOIseq method were overlapped with that from PossionDis method (**Table 3**).
Furthermore, nine DEGs were selected to validate the reliability of sequencing data,
and the results indicated that the expression trends of these DEGs were consistent
with those obtained from transcriptome data (**Figure 3**), and genes name, genes ID
and primers information were shown in **Table S1**. Moreover, correction analysis
results showed that the fold change values from the two methods were highly
correlated (Pearson correlation coefficient $R = 0.87$) at a high level of statistical
significance ($P < 0.01$). These results suggest that reliable sequencing data are obtained
from genome-wide transcriptome.

QTL location analysis of DEGs

To obtain potential candidate genes affecting GP, we conducted a mapping
analysis of DEGs from PossionDis method with the QTL location information of all
traits in the pig QTL database. Furthermore, the DEGs located in GP related QTLs,

including average glycogen (GLY), average glycolytic potential (GLYPO), average lactate, Glucose-6-phosphate content (GLU6P) and residual glycogen (RESGLY) were refined. In total, 97 non-redundant DEGs were mapped to GP related QTLs from three paired comparisons (H11vs L11, H9 vs L9, H5 vs L5). Moreover, 41 non-redundant DEGs from H11vs L11, 61 non-redundant DEGs from H9 vs L9, and 12 non-redundant DEGs from H5 vs L5 located in GP related QTLs. The detailed QTL mapping information of DEGs from each paired group was shown in **File S7**.

Identification of potential differential and specific SNPs and Indels

In the present study, six higher GP pigs and six lower GP pigs were used for sequencing analysis. Therefore, another objective of this study was to find potential differential and specific SNPs and Indels between GP higher and lower groups. SNP variation types and numbers were analyzed through comparison with reference genome. About 70,000 SNPs were identified in each sequencing library, and transition is the main SNP variation type (**File S1-1**).The detailed SNP annotation information for each sequencing library, including chromosome, position, gene region,

and gene ID were shown in **File S1-2** to **File S1-7**. Subsequently, the potential differentially and specifically SNPs between H and L GP groups were refined. A total of 86,381 potential differentially SNPs (**File S1-8**), and 1,076 potential specific SNPs were figured out (**File S1-9**). Notably, 40 DEGs derived from PossionDis method contained potential specific SNPs were screened (**Table 4**), while only one DEG from NOIseq method, named C-X-C motif chemokine ligand 10 or C-X-C motif chemokine 10 precursor (CXCL10, ENSSSCG00000008977), contained a potential specific SNP (Chromosome 8, Position 75830488). Furthermore, eight potential differentially SNPs were validated using PCR product sequencing method. The results showed that six of eight were real SNPs, the genotype and allele frequencies of each variation in higher and lower GP groups were shown in **Table 5**.

Additionally, about 5,000-7,000 Indels were identified from each sequencing library when compared with reference genome (**File S8-1 to File S8-6**), and 292,162 and 66 potential differential Indels were figured out from H11 vs L11 (**File S8-7**), H9 vs L9 (**File S8-8**), and H5 vs L5 (**File S8-9**) groups, respectively. Moreover, the number of Indels was reduced as the length increase, **File S9** showed the annotation

information of more than 5bp length indels.

DISCUSSION

Glycogen metabolism disorder involved in defects in glycogen metabolism or excess glycogen stored in liver and muscle cause glycogen storage diseases. In livestock, measurement of muscle glycogen (G) combined with measure of glucose (MG), glucose-6-phosphate (G6P) and lactate content called glycolytic potential (GP) to evaluate the capability of glycogen storage and using it to evaluate the meat quality development. Therefore, we measured the GP phenotypic values of each pig derived from a 279 P×D×L×Y commercial pigs and selected the pigs with extreme GP phenotypic values for RNA-seq analysis in this study. In order to meet the requirements of biological replicates and use as many experimental pigs as possible in sequencing, we selected twelve experimental pigs from three slaughtered batches (No.11, 9 and 5) and four individuals with extreme GP phenotypic values for each batch (two with the higher GP and two with the lower GP) to construct RNA sequencing libraries (**Figure S1**). Factually, the phenotypic values of many other meat quality traits showed obvious differences between the higher and lower GP groups,

and several reached statistical significant differences, including pH_{45min}, DL_{48h}, and CL (**Table 1**). Previous studies indicated that GP was closely related to pH value, drip loss and cooking loss (Hamilton et al. 2003; Zybert et al. 2013), which suggest that the pigs selected are suitable for sequencing analysis.

In the present study, six sequencing libraries were constructed and high-quality sequencing data were obtained through quality-control analysis, including Q20 and Q30, reads mapping, and base composition analysis. The ratio of clean reads with > Q20 and > Q30 reached 97.0% and 93.0% for each library, respectively. The ratio of mapped reads reached more than 65.0% for each library, which is consisted with previous studies (Chen et al. 2011; Esteve-Codina et al. 2011; Li et al. 2016b), but less than the some studies where more than 70.0% of reads were mappable (Li et al. 2016a; Ramayo-Caldas et al. 2012). These discrepancies may be caused by the differences of reference genome version and alignment software. Generally, the reads obtained from six libraries can meet the subsequent analysis of differentially expressed genes identification and variations detection.

One of the main objectives of this study to identify the candidate genes affecting

GP related traits. Based on the strategy of RNA-seq libraries construction (**Figure S1**), we adopted two methods to screen the DEGs in the present study. PossionDis method was used to identify the DEGs considering paired comparisons of H11 vs L11, H9 vs L9, and H5 vs L5. Satisfactorily, hundreds of DEGs were found in paired comparative analysis and a certain number of DEGs were overlapped between two groups, although only six DEGs were overlapped among three groups (**Figure 2**). In RNA-seq analysis, some questions were still difficulty to avoid in livestock, such as the large variations were often happened between samples. These phenomenon was also observed in this study, which may be the reason that only a few overlapping genes were identified among each RNA-sequencing libraries. Moreover, the $FDR \leq 0.001$ and $\log_2 \text{Foldchange} \geq 1$ were considered to screen the DEGs in this study, so some truly DEGs may be ignored. Moreover, NOIseq method was applied to identify the DEGs between H and L groups considering biological replicates, 69 DEGs were detected in this study. Notably, 36 DEGs from NOIseq method were overlapped that from PossionDis method (**Table 3**). Although the major genes related to GP have not been confirmed, the DEGs found in this study pave good foundation for identifying

the genuine major genes affecting GP.

At present, some QTLs related to average glycogen (18), average glycolytic potential (20), average lactate (10), glucose-6-phosphate content (3) and residual glycogen (6) were deposited in pig QTLs database, but it is still difficult to confirm the functional genes affecting glycogen metabolism due to the large size of these QTLs. To refine the candidate genes related to glycogen metabolism, we performed an integrated analysis between porcine QTLs and DEGs according the genomics location information. Fortunately, we found 97 potential candidate genes affecting GP-related traits (**File S7**), although the functions of these candidate genes are still needed to confirm. GP respond to the glycogen storage capability, which can be used to evaluate the meat quality development, and is closely related to several meat quality traits. Up to date, only two major genes and corresponding cause mutations affecting GP have been confirmed, one is *PRKAG3*(Milan et al. 2000) and another is *PHKG1* (Ma et al. 2014). Notably, harmful allele (*RN* allele) in *PRKAG3* gene is exclusively found purebred Hampshire breed and its related synthetic lines, and the effect is only observed in purebred Hampshire breed, which limited its application in

most pig breeding programs. *PHKG1* is the second major gene related to glycogen metabolism and g.8283C>A is the cause mutation, while the mutation is mainly observed in Duroc breeds, especially in White Duroc breed (Ma et al. 2014), which can be implicated for screening of excellent Duroc pigs, but the application value in other pig breeds are limited. Therefore, additional major genes and mutations are needed to be explored for GP-related traits molecular breeding in most of pig breeds. In the present study, the experimental pigs are from a four-hybrid commercial pig (P×D) × (L×Y) enriched in genetic diversities, which suggest the candidate genes and variations identified in this study might be applied more widely.

Recent years, the genome sequences and high-density SNP chips are widely used for detection of variations significantly associated to economic traits in livestock (Liu et al. 2015; Ma et al. 2014; Rubin et al. 2012; Rubin et al. 2010; Verardo et al. 2016). In this study, we found 1,076 potential specific SNPs between the higher and the lower GP group (**File S1-9**) using the transcriptome sequencing data, which are the potential variations associated to GP-related traits. Notably, 40 DEGs contains potential specific SNPs (**Table 4**), and most of SNPs are located in gene coding

sequence (CDS) which is in accordance with the data from transcriptome sequencing.

Factually, many causal mutations of major genes identified in previous studies are in

CDS region and cause the change of amino acid, such as the g.1843C>T mutation in

RYR1 gene (Fujii et al. 1991), the R200Q mutation in *PRKAG3* gene (Milan et al.

2000), the G>A mutation in CDS of *MC4R* gene (Kim et al. 2000). Therefore, the

potential specific SNPs identified in this study can provide valuable references for

identifying the functional mutations associated with GP-related traits. In the present

study, a potential specific SNP was included in *CXCL10* gene, while previous study

reported that *CXCL10* was closely related to Type1 diabetes (Antonelli et al. 2014).

Intriguingly, a SNP was validated in exon of differentially expressed gene *AGL* in this

study. *AGL*, a key glycogen debranching enzyme, plays an important role in reducing

glycogen degradation and accumulating of limit dextrin in skeletal muscle. It was

previously reported that a causal mutation in human *AGL* exon has close relationship

with glycogen storage disease type III (Rousseau-Nepton et al. 2015).

Moreover, the indels were analyzed and hundreds of potential differentially

Indels were figured out between the higher and lower GP groups. To our special

attention, the effects of indels are often enormous, for example, an 11bp deletion in the third exon of *myostatin* gene (*GDF-8*) causes the double-muscling phenotype in Belgian Blue cattle (Grobet et al. 1997; McPherron and Lee 1997). Therefore, although only a few number of large insertion and deletion are observed in our data, approximately 40 Indels with a length ≥ 5 bp were identified in each sequencing library (**File S8**), however, the relationship between these indels and GP-related traits, and regulatory mechanisms might become the focus in future study.

CONCLUSIONS

In livestock, GP, responding to the glycogen storage capability, is a critical indicator for evaluating meat quality. In the present study, we conducted high-throughput transcriptome sequencing to identify the potential candidate genes and variations related to GP-related traits using the muscle samples from a four-hybrid commercial pig (P×D×L×Y). A total of 525, 698 and 135 DEGs were identified between H11 vs L11, H9 vs L9, and H5 vs L5 GP groups using PoissonDis method, respectively. While, a total of 69 DEGs were identified between H (H11, H9 and H5)

and L (L11, L9 and L5) GP groups using NOIseq method. Moreover, 1,076 potential specific SNPs were figured out between H and L GP groups, and approximately 40 large Indels with a length ≥ 5 bp were identified in each sequencing library. Taken together, our data provide foundation for further identifying the key genes and screening the functional mutations affecting GP-related traits.

ACKNOWLEDGEMENTS

We thank Jiaolong Li from Nanjing Agricultural University for his help in samples collection. This work was supported by the National Natural Science Foundation of China (31501920), the Fundamental Research Funds of the Central Universities (KJQN201605), the National Major Project of Breeding for Transgenic Pig (2016ZX08006001-003) and the Program of Hainan Provincial Engineering Research Center for Livestock and Poultry Cultivation (XQ2016JSZX001).

445

446 LITERATURE CITED

- 447 Antonelli, A, Ferrari, SM, Corrado, A, Ferrannini, E, and Fallahi, P (2014) CXCR3,
448 CXCL10 and type 1 diabetes. *Cytokine Growth F R* 25: 57-65.
- 449 Audic, S, and Claverie, JM (1997) The significance of digital gene expression profiles.
450 *Genome Res* 7: 986-995.
- 451 Bee, G, Anderson, AL, Lonergan, SM, and Huff-Lonergan, E (2007) Rate and extent
452 of pH decline affect proteolysis of cytoskeletal proteins and water-holding capacity in
453 pork. *Meat science* 76: 359-365.
- 454 Chen, CY, Ai, HS, Ren, J, Li, WB, Li, PH, Qiao, RM, Ouyang, J, Yang, M, Ma, JW,
455 and Huang, LS (2011) A global view of porcine transcriptome in three tissues from a
456 full-sib pair with extreme phenotypes in growth and fat deposition by paired-end
457 RNA sequencing. *Bmc Genomics* 12.
- 458 Esteve-Codina, A, Kofler, R, Palmieri, N, Bussotti, G, Notredame, C, and
459 Perez-Enciso, M (2011) Exploring the gonad transcriptome of two extreme male pigs
460 with RNA-seq. *Bmc Genomics* 12.
- 461 Fujii, J, Otsu, K, Zorzato, F, Deleon, S, Khanna, VK, Weiler, JE, Obrien, PJ, and
462 MacLennan, DH (1991) Identification of a Mutation in Porcine Ryanodine Receptor
463 Associated with Malignant Hyperthermia. *Science* 253: 448-451.
- 464 Grobet, L, Martin, LJR, Poncelet, D, Pirottin, D, Brouwers, B, Riquet, J, Schoeberlein,
465 A, Dunner, S, Menissier, F, Massabanda, J, *et al.* (1997) A deletion in the bovine
466 myostatin gene causes the double-muscled phenotype in cattle. *Nat Genet* 17: 71-74.
- 467 Hamilton, DN, Miller, KD, Ellis, M, McKeith, FK, and Wilson, ER (2003)
468 Relationships between longissimus glycolytic potential and swine growth
469 performance, carcass traits, and pork quality. *J Anim Sci* 81: 2206-2212.
- 470 Jasinska, K, and Kurek, MA (2017) The effect of oil plants supplementation in pig
471 diet on quality and nutritive value of pork meat. *Anim Sci Pap Rep* 35: 137-146.
- 472 Kaminski, S, Kocwin-Podsiadla, M, Sieczkowska, H, Help, H, Zybert, A, Krzecio, E,
473 Antosik, K, Brym, P, Wojcik, E, and Adamczyk, G (2010) Screening 52 single

nucleotide polymorphisms for extreme value of glycolytic potential and drip loss in
pigs. *J Anim Breed Genet* 127: 125-132.

Kim, D, Landmead, B, and Salzberg, SL (2015) HISAT: a fast spliced aligner with
low memory requirements. *Nat Methods* 12: 357-U121.

Kim, KS, Larsen, N, Short, T, Plastow, G, and Rothschild, MF (2000) A missense
variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness,
growth, and feed intake traits. *Mamm Genome* 11: 131-135.

Langmead, B, and Salzberg, SL (2012) Fast gapped-read alignment with Bowtie 2.
Nat Methods 9: 357-U354.

Lebret, B (2008) Effects of feeding and rearing systems on growth, carcass
composition and meat quality in pigs. *Animal : an international journal of animal
bioscience* 2: 1548-1558.

Li, B, and Dewey, CN (2011) RSEM: accurate transcript quantification from
RNA-Seq data with or without a reference genome. *Bmc Bioinformatics* 12.

Li, BJ, Dong, C, Li, PH, Ren, ZQ, Wang, H, Yu, FX, Ning, CB, Liu, KQ, Wei, W,
Huang, RH, *et al.* (2016a) Identification of candidate genes associated with porcine
meat color traits by genome-wide transcriptome analysis. *Sci Rep-Uk* 6.

Li, BJ, Liu, KQ, Weng, QN, Li, PH, Wei, W, Li, QF, Chen, J, Huang, RH, Wu, WJ,
and Liu, HL (2016b) RNA-seq analysis reveals new candidate genes for drip loss in a
Pietrain x Duroc x Landrace x Yorkshire population. *Anim Genet* 47: 192-199.

Liu, XX, Xiong, XW, Yang, J, Zhou, LS, Yang, B, Ai, HS, Ma, HB, Xie, XH, Huang,
YX, Fang, SM, *et al.* (2015) Genome-wide association analyses for meat quality traits
in Chinese Erhualian pigs and a Western Duroc x (Landrace x Yorkshire) commercial
population. *Genet Sel Evol* 47.

Ma, JW, Yang, J, Zhou, LS, Ren, J, Liu, XX, Zhang, H, Yang, B, Zhang, ZY, Ma, HB,
Xie, XH, *et al.* (2014) A Splice Mutation in the PHKG1 Gene Causes High Glycogen
Content and Low Meat Quality in Pig Skeletal Muscle. *Plos Genet* 10.

McKenna, A, Hanna, M, Banks, E, Sivachenko, A, Cibulskis, K, Kernytsky, A,
Garimella, K, Altshuler, D, Gabriel, S, Daly, M, *et al.* (2010) The Genome Analysis
Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing
data. *Genome Res* 20: 1297-1303.

McPherron, AC, and Lee, SJ (1997) Double muscling in cattle due to mutations in the

myostatin gene. *P Natl Acad Sci USA* 94: 12457-12461.

Milan, D, Jeon, JT, Looft, C, Amarger, V, Robic, A, Thelander, M, Rogel-Gaillard, C, Paul, S, Iannuccelli, N, Rask, L, *et al.* (2000) A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. *Science* 288: 1248-1251.

Monin, G, and Sellier, P (1985) Pork of low technological quality with a normal rate of muscle pH fall in the immediate post-mortem period: The case of the Hampshire breed. *Meat science* 13: 49-63.

Pertea, M, Pertea, GM, Antonescu, CM, Chang, TC, Mendell, JT, and Salzberg, SL (2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 33: 290-+.

Popp, J, Wicke, M, Klein, G, and Krischek, C (2015) The relationship of pork longissimus muscle pH to mitochondrial respiratory activities, meat quality and muscle structure. *Animal : an international journal of animal bioscience* 9: 356-361.

Pugliese, C, and Sirtori, F (2012) Quality of meat and meat products produced from southern European pig breeds. *Meat science* 90: 511-518.

Ramayo-Caldas, Y, Mach, N, Esteve-Codina, A, Corominas, J, Castello, A, Ballester, M, Estelle, J, Ibanez-Escriche, N, Fernandez, AI, Perez-Enciso, M, *et al.* (2012) Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. *Bmc Genomics* 13.

Rousseau-Nepton, I, Okubo, M, Grabs, R, Mitchell, J, Polychronakos, C, Rodd, C, and Consortium, FC (2015) A founder AGL mutation causing glycogen storage disease type IIIa in Inuit identified through whole-exome sequencing: a case series. *Can Med Assoc J* 187: E68-E73.

Rubin, CJ, Megens, HJ, Barrio, AM, Maqbool, K, Sayyab, S, Schwochow, D, Wang, C, Carlborg, O, Jern, P, Jorgensen, CB, *et al.* (2012) Strong signatures of selection in the domestic pig genome. *P Natl Acad Sci USA* 109: 19529-19536.

Rubin, CJ, Zody, MC, Eriksson, J, Meadows, JRS, Sherwood, E, Webster, MT, Jiang, L, Ingman, M, Sharpe, T, Ka, S, *et al.* (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464: 587-U145.

Shen, SH, Park, JW, Lu, ZX, Lin, L, Henry, MD, Wu, YN, Zhou, Q, and Xing, Y (2014) rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *P Natl Acad Sci USA* 111: E5593-E5601.

538 Sieczkowska, H, Zybert, A, Krzecio, E, Antosik, K, Kocwin-Podsiadla, M, Pierzchala,
539 M, and Urbanski, P (2010) The expression of genes PKM2 and CAST in the muscle
540 tissue of pigs differentiated by glycolytic potential and drip loss, with reference to the
541 genetic group. *Meat science* 84: 137-142.

542 Tarazona, S, Garcia-Alcalde, F, Dopazo, J, Ferrer, A, and Conesa, A (2011)
543 Differential expression in RNA-seq: A matter of depth. *Genome Res* 21: 2213-2223.

544 Trapnell, C, Roberts, A, Goff, L, Pertea, G, Kim, D, Kelley, DR, Pimentel, H,
545 Salzberg, SL, Rinn, JL, and Pachter, L (2012) Differential gene and transcript
546 expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature*
547 *protocols* 7: 562-578.

548 Verardo, LL, Silva, FF, Lopes, MS, Madsen, O, Bastiaansen, JWM, Knol, EF, Kelly,
549 M, Varona, L, Lopes, PS, and Guimaraes, SEF (2016) Revealing new candidate genes
550 for reproductive traits in pigs: combining Bayesian GWAS and functional pathways.
551 *Genet Sel Evol* 48.

552 Werner, C, Natter, R, Schellander, K, and Wicke, M (2010) Mitochondrial respiratory
553 activity in porcine longissimus muscle fibers of different pig genetics in relation to
554 their meat quality. *Meat science* 85: 127-133.

555 Zybert, A, Sieczkowska, H, Antosik, K, Krzecio-Nieczyporuk, E, Adamczyk, G, and
556 Kocwin-Podsiadla, M (2013) Relationship between Glycolytic Potential and Meat
557 Quality of Duroc Pigs with Consideration of Carcass Chilling System. *Ann Anim Sci*
558 13: 645-654.

559 **Table 1** Statistical analysis for experimental pigs with extreme phenotypic GP value

Groups ^a	H5	L5		H9	L9		H11	L11		H	L	
	N=2	N=2		N=2	N=2		N=2	N=2		N=6	N=6	
Traits ^b	Mean ± SD ^c		P-value	Mean ± SD		P-value	Mean ± SD		P-value	Mean ± SD		P-value
GP (μmol/g)	172.91±1.24	114.44±0.16	0.0002	164.59±0.33	133.37±3.71	0.0070	195.78±21.70	140.32±7.37	0.0758	177.76±17.41	129.38±12.54	0.0002
RG (μmol/g)	12.41±0.45	7.96±1.30	0.0447	10.55±0.39	7.24±1.30	0.0750	11.42±1.64	7.82±0.74	0.1053	11.46±1.14	7.67±0.95	0.0001
MG (μmol/g)	6.17±0.99	3.96±0.47	0.1053	7.19±0.16	4.37±2.38	0.2362	8.56±2.09	4.39±1.51	0.1495	7.30±1.49	4.24±1.29	0.0035
G (μmol/g)	6.24±0.55	4.00±0.83	0.0869	3.35±0.54	2.88±1.08	0.6357	2.87±0.45	3.43±0.77	0.4675	4.15±1.68	3.44±0.86	0.3742
G6P (μmol/g)	1.31±0.02	1.31±0.01	0.8076	1.58±0.07	1.56±0.13	0.8375	1.46±0.15	1.68±0.198	0.3273	1.45±0.14	1.52±0.19	0.5282
LA (μmol/g)	145.47±0.30	95.91±2.72	0.0015	140.34±0.30	115.78±6.05	0.0290	170.04±18.12	121.33±8.46	0.0749	151.95±16.35	111.01±12.88	0.0007
pH _{45min}	5.95±0.34	6.25±0.26	0.4227	6.07±0.50	6.46±0.18	0.4057	6.19±0.16	6.61±0.01	0.0637	6.07±0.299	6.44±0.21	0.0333
pH _{24h}	5.54±0.05	5.70±0.05	0.0838	5.59±0.01	5.58±0.01	0.3117	5.49±0.01	5.55±0.13	0.5457	5.54±0.05	5.61±0.09	0.1377
DL _{24h} (%)	1.85±0.71	1.31±0.15	0.3978	3.56±2.67	1.56±0.18	0.3912	1.42±0.04	1.31±0.13	0.3777	2.27±1.59	1.37±0.16	0.1990
DL _{48h} (%)	4.48±0.67	2.64±0.34	0.0733	5.63±3.06	2.85±0.12	0.3265	4.01±0.24	2.59±0.403	0.0501	4.70±1.59	2.69±0.27	0.0119
CL (%)	32.70±0.92	30.72±1.87	0.3107	34.33±0.99	31.01±1.46	0.1170	31.56±4.14	24.68±0.45	0.1451	32.86±2.32	28.80±3.37	0.0356
CW (kg)	76.30±0.99	86.90±4.39	0.0793	77.30±3.82	89.20±7.64	0.1875	86.10±16.26	77.70±4.95	0.5570	79.90±8.90	84.60±7.07	0.3352

560 ^a Groups: H represents the pigs with extreme high GP value; L represents the pigs with extreme lower GP value; N represents the number of experimental pigs in each group.

561 ^b Traits: CW: Carcass weight; DL_{24h}: Drip loss measured at 24h postmortem; DL_{48h}: Drip loss measured at 48h postmortem; CL: Cooking loss; MG: Muscle glycogen; G:
562 Glucose; G6P: Glucose-6-phosphate; LA: Lactic acid; GP: Glycolytic potential; RG: Residual glycogen.

563 ^c SD: standard deviation.

Table 2 Output statistics and annotation information of sequencing reads

Samples	H11	H9	H5	L11	L9	L5
Total Raw Reads	52261610	53893752	52261668	52261602	53893960	50628480
Total Clean Reads	44993154	45347028	45385424	45136800	45134096	44466302
Total Clean Reads Ratio (%)	86.09	84.14	86.84	86.37	83.75	87.83
Total Adatper Reads	6210206	7439632	5912532	5982576	7761344	5190604
Total Adatper Reads Ratio (%)	11.88	13.80	11.31	11.45	14.40	10.25
Total High Unknown Base Reads	21968	22364	10256	21902	21254	14352
Total High Unknown Base Reads Ratio (%)	0.04	0.04	0.02	0.04	0.04	0.03
Total Low Quality Reads	1036282	1084728	953456	1120324	977266	957222
Total Low Quality Reads Ratio (%)	1.98	2.01	1.82	2.14	0.18	1.89
Clean reads Q20 (%)	97.73	97.71	97.72	97.64	97.66	97.63
Clean reads Q30 (%)	93.62	93.58	93.85	93.44	93.43	93.6
Total mapped reads	29629518	29244540	29843848	29824708	29420544	29634178
Total_mapped_reads ratio (%)	65.85	64.49	65.76	66.08	65.18	66.64
Unique_mapped_reads	26859358	26362290	26904222	26903986	26632142	26760830
Unique_mapped_reads ratio (%)	59.70	58.13	59.28	59.61	59.01	60.18
Multiple mapped reads	2770160	2882250	2939626	2920722	2788402	2873348
Multiple mapped reads ratio (%)	6.16	6.36	6.48	6.47	6.18	6.46
Splice mapped reads unmapped	15363636	16102488	15541576	15312092	15713552	14832124
Splice mapped reads ratio (%)	34.15	35.51	34.24	33.92	34.82	33.36
Total Gene Number	14768	13841	13976	14201	14373	14058
Known Gene Number	14067	13187	13312	13516	13687	13394
Novel Gene Number	701	654	664	685	686	664
Total Transcript Number	20554	19069	19210	19664	19959	19345
Known Transcript Number	13385	12221	12310	12600	12846	12423
Novel Transcript Number	7169	6848	6900	7064	7113	6922

565 **Table 3** Overlapped DEGs from NOIseq method compared with PossionDis method

Overlapped genes ID	Overlapped genes name	Length	H-Expression	L-Expression	log2FoldChange (L/H)	Probability	Up/Down-Regulation (L/H)
ENSSSCG00000030999	<i>MYH6</i>	3402	23.506667	4.6033333	-2.352319	0.8335427	Down
ENSSSCG00000028326	<i>Novel gene</i>	1127	26.793333	55.676667	1.0551988	0.7304825	Up
ENSSSCG00000028235	<i>Novel gene</i>	1532	25.926667	0.01	-11.34022	0.9766946	Down
ENSSSCG00000028056	<i>Novel gene</i>	1791	2.24	6.6466667	1.5691323	0.704386	Up
ENSSSCG00000027982	<i>Novel gene</i>	813	6.78	20.236667	1.5776145	0.7769338	Up
ENSSSCG00000025968	<i>Novel gene</i>	2272	0.01	2.0133333	7.6534422	0.7775419	Up
ENSSSCG00000024953	<i>Novel gene</i>	688	2.9233333	8.5833333	1.5539237	0.7226974	Up
ENSSSCG00000023783	<i>Novel gene</i>	653	7.06	0.01	-9.463524	0.9255781	Down
ENSSSCG00000023287	<i>Novel gene</i>	964	158.14667	73.143333	-1.112465	0.7492723	Down
ENSSSCG00000018006	<i>Novel gene</i>	309	135.63333	284.84667	1.0704737	0.7459729	Up
ENSSSCG00000017723	<i>CCL2</i>	863	3.2166667	14.376667	2.1600909	0.8050439	Up

566 **Table 4** DEGs PossionDis method containing potential specific SNPs

Chromosome	Genome Position	Gene Region	Ensemble Gene ID	Ensemble Gene Name	Length
1	144542379	CDS	ENSSSCG00000004738	<i>Novel gene</i>	6890
1	180980461	CDS	ENSSSCG00000004931	<i>IGDCC4</i>	4662
2	69465021	gene	ENSSSCG000000013655	<i>ICAM1</i>	3083
2	90534227	CDS	ENSSSCG000000014117	<i>THBS4</i>	3095
2	117842636	gene	ENSSSCG000000014190	<i>Novel gene</i>	5117
2	90219536	CDS	ENSSSCG000000026098	<i>CMYA5</i>	11424
2	61525703	CDS	ENSSSCG000000027903	<i>Novel gene</i>	1691
3	293372	CDS	ENSSSCG000000007543	<i>Novel gene</i>	2515
3	877756	gene	ENSSSCG000000007554	<i>ZFAND2A</i>	3211
3	53962701	CDS	ENSSSCG000000008153	<i>SLC9A2</i>	2636
4	47616616	CDS	ENSSSCG000000006114	<i>CU207250.1</i>	1018
4	97119619	CDS	ENSSSCG000000006359	<i>ADAMTS4</i>	3299
4	104971015	CDS	ENSSSCG000000006582	<i>S100A14</i>	1102
4	129775973	CDS	ENSSSCG000000006872	<i>AGL</i>	4328
4	135027605	CDS	ENSSSCG000000006891	<i>GCLM</i>	1613

4	139663858	CDS	ENSSSCG00000006924	<i>GBP1</i>	3155
5	14287316	exon	ENSSSCG000000027055	<i>NUAK1</i>	5756
7	28064666	CDS	ENSSSCG00000001440	<i>NOTCH4</i>	6727
7	33595838	gene	ENSSSCG00000001499	<i>DST</i>	23654
7	47070440	exon	ENSSSCG00000001711	<i>Novel gene</i>	3122
7	129349501	CDS	ENSSSCG00000002528	<i>Novel gene</i>	2660
7	27923781	CDS	ENSSSCG000000025712	<i>Novel gene</i>	4229
8	34336166	gene	ENSSSCG000000008799	<i>Novel gene</i>	7380
8	75830488	exon	ENSSSCG000000008977	<i>CXCL10</i>	1258
8	89412299	exon	ENSSSCG000000009048	<i>GAB1</i>	4152
8	140235738	exon	ENSSSCG000000009217	<i>PKD2</i>	3979
8	5946666	CDS	ENSSSCG000000029711	<i>ZNF518B</i>	3243
9	9177239	gene	ENSSSCG000000014834	<i>UCP3</i>	3828
9	143700687	gene	ENSSSCG000000015595	<i>ATF3</i>	1976
9	10407273	exon	ENSSSCG000000021638	<i>NEU3</i>	6093
12	13577371	CDS	ENSSSCG000000017268	<i>PRKCA</i>	1480
12	58218924	CDS	ENSSSCG000000018003	<i>MYH2</i>	5893
13	38190855	gene	ENSSSCG000000011452	<i>Novel gene</i>	1172

13	76193324	CDS	ENSSSCG000000011592	<i>PLXND1</i>	7197
13	134683501	gene	ENSSSCG000000011810	<i>Novel gene</i>	3659
13	73648039	gene	ENSSSCG000000023084	<i>Novel gene</i>	3112
14	71902350	CDS	ENSSSCG000000010226	<i>JMJD1C</i>	4960
14	24766922	CDS	ENSSSCG000000028018	<i>EP400</i>	8484
17	57474885	exon	ENSSSCG000000007463	<i>PTGIS</i>	5380
GL896332.1	25236	CDS	ENSSSCG000000029168	<i>MARF1</i>	4505

567 **Table 5** Validation of the potential differentially SNPs in higher and lower GP groups

<i>AGL gene</i> (Gene ID: 100156940)		Chromosome: 4	Position: 129775973	Variation: T/C		
		Genotype			Allele frequency	
	Number	C/C	C/T	T/T	C	T
H group	6	0	0	6	0.00	1.00
L group	6	0	4	2	0.33	0.67

<i>MAPKB gene</i> (Gene ID: 100158174)		Chromosome: 1	Position: 144542379	Variation: T/C		
		Genotype			Allele frequency	
	Number	C/C	C/T	T/T	C	T
H group	6	1	3	2	0.42	0.58
L group	6	3	3	0	0.75	0.25

<i>CMYA5 gene</i> (Gene ID: 100505410)		Chromosome: 2	Position: 90219536	Variation: T/C		
		Genotype			Allele frequency	
	Number	C/C	C/T	T/T	C	T
H group	6	6	0	0	1.00	0.00
L group	6	6	0	0	1.00	0.00

<i>NEU3 gene</i> (Gene ID: 100739223)		Chromosome: 9	Position: 10407273	Variation: T/C		
		Genotype			Allele frequency	
	Number	C/C	C/T	T/T	C	T
H group	6	0	0	6	0.00	1.00
L group	6	0	0	6	0.00	1.00

<i>NUAK1 gene</i>		Chromosome:	Position:	Variation:		
(Gene ID: 100523669)		5	14287316	T/C		
Genotype				Allele frequency		
	Number	C/C	C/T	T/T	C	T
H group	6	1	4	1	0.50	0.50
L group	6	1	3	2	0.42	0.58

<i>PRKCA gene</i>		Chromosome:	Position:	Variation:		
(Gene ID: 397184)		12	13577371	T/C		
Genotype				Allele frequency		
	Number	C/C	C/T	T/T	C	T
H group	6	5	1	0	0.92	0.08
L group	6	3	3	0	0.75	0.25

<i>PTGIS gene</i>		Chromosome:	Position:	Variation:		
(Gene ID: 100126284)		17	57474885	T/C		
Genotype				Allele frequency		
	Number	C/C	C/T	T/T	C	T
H group	6	2	4	0	0.67	0.33
L group	6	5	1	0	0.92	0.08

<i>UCP3 gene</i>		Chromosome:	Position:	Variation:		
(Gene ID: 397116)		9	9177239	T/C		
Genotype				Allele frequency		
	Number	C/C	C/T	T/T	C	T
H group	6	0	2	4	0.83	0.17
L group	6	1	4	1	0.50	0.50

569

570 **Figure 1** Volcano-plot and statistic of DEGs from PossionDis method

571 (a) Red and blue color represents significantly up-regulated and down-regulated genes,
572 respectively. (b) H11-vs-L11, H9-vs-L9 and H5-vs-L5 indicate the comparative
573 strategy by PossionDis method in X axis. Y axis indicates the gene number.

574 **Figure 2** Venn of DEGs obtained from PossionDis method

575 Overlapping analysis of DEGs was performed using web tool
576 (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). Different color represents
577 different combination, and the number in the overlap region represents the overlapped
578 DEGs number. In total, 27 DEGs were overlapped between H11 vs L11 (HL11) and
579 H5 vs L5 (HL5), 26 DEGs were overlapped between H9 vs L9 (HL9) and H5 vs L5
580 (HL5), 157 DEGs were overlapped between H11 vs L11 (HL11) and H9 vs L9 (HL9),
581 and 6 DEGs were overlapped among three groups.

582 **Figure 3** Validation of DEGs by real-time PCR

583 Total RNA of twelve LD muscle samples were extracted, which derived from three
584 slaughtered batches (No. 5, 9 and 11), four individuals for each batch (two with the

higher GP and two with the lower GP). FPKM values are used to calculate the gene expression in RNA-seq and normalize the expression of one group to “1”. In real-time PCR, relative expression levels are calculated using $\Delta\Delta C_t$ value method and normalized by reference gene *HPRT*, and similarly normalize the expression of one group to “1”. The paired Student’s *t*-test is used to evaluate the statistical significance of differences between the two groups, * $P \leq 0.05$, ** $P \leq 0.01$. All data are presented as mean $\pm$ standard error (SE).

Figure S1 The schematic overview of RNA-seq libraries construction.

Figure S2 Length distribution of clean reads. Y axis represents the number of total clean reads with certain length. X axis represents the length of clean reads

ADM

EGR1

GLUT3

NOR-1

THBS1

ADAMTS8

CITED2

GCAT

RASSF4

H11

L11

H9

L9

H5

L5