

引文格式: 荣丽云, 黄宁, 王孝义, 等. 基于 WGCNA 方法识别影响猪肉品质性状的关键模块和基因[J]. 云南农业大学学报(自然科学), 2022, 37(5): 769–779. DOI: 10.12101/j.issn.1004-390X(n).202205013

基于 WGCNA 方法识别影响猪肉品质性状的关键模块和基因*

荣丽云, 黄宁, 王孝义, 李明丽, 陈强**, 鲁绍雄**
(云南农业大学 动物科学技术学院, 云南 昆明 650201)

摘要: 【目的】发掘和识别猪肉品质性状相关的基因共表达模块和关键基因。【方法】采用加权基因共表达网络分析 (weighted gene co-expression network analysis, WGCNA)、基因互作关系网络以及中心性方法, 对来自 NCBI 基因表达数据库的大蒲莲猪和长白猪各 8 头个体的背最长肌基因表达数据集 GSE119368 进行关键基因共表达模块和基因挖掘, 以 3 头撒坝猪和 3 头大白猪为素材对肉质性状部分关键基因进行 qPCR 验证。【结果】共鉴定出蓝色和红色 2 个与猪肉品质性状显著相关的基因共表达模块 (r 值分别为 0.89 和 0.91, P 值分别为 $5E-6$ 和 $7E-7$); 蓝色模块基因主要与快慢肌纤维转换及细胞外基质等相关, 其中的 10 个基因 *COL3A1*、*COL5A1*、*COL1A2*、*COL14A1*、*CRTAP*、*COL5A3*、*LEPREL1*、*DEPTOR*、*COL4A2* 和 *COL6A2* 为肉质性状关键基因; 红色模块基因主要与骨骼肌细胞及脂肪细胞分化等有关, 其中的 11 个基因 *FBXL4*、*RNF115*、*FBXO40*、*ASB4*、*ASB15*、*UBE2L6*、*RNF41*、*ASB5*、*HERC5*、*SOC1* 和 *ASB11* 为肉质性状关键基因; 蓝色模块的前 3 个关键基因在撒坝猪和大白猪背最长肌中表达差异极显著 ($P < 0.01$), 红色模块的前 3 个基因表达差异不显著 ($P > 0.05$)。【结论】研究结果初步鉴定了不同肉质类型猪种间肉质相关的基因共表达模块和关键基因, 但相同肉质类型猪种的肉质性状关键基因及其效应可能存在明显差异。

关键词: 猪; 肉质性状; 基因表达; 加权基因共表达网络; 关键基因

中图分类号: S828.91 文献标志码: A 文章编号: 1004-390X (2022) 05-0769-11

Identification of Key Gene Co-expression Modules and Genes Associated with Pork Quality Based on WGCNA Method

RONG Liyun, HUANG Ning, WANG Xiaoyi, LI Mingli, CHEN Qiang, LU Shaoxiong
(Faculty of Animal Science and Technology, Yunnan Agricultural University, Kunming 650201, China)

Abstract: [Purpose] To explore and identify the key gene co-expression modules and genes associated with meat quality trait in pigs. [Methods] The key gene co-expression modules and genes were investigated using gene expression data GSE119368 of longissimus dorsi muscles including eight Dapulian and eight Landrace pigs from NCBI GEO database by weighted gene co-expression network analysis (WGCNA), gene interaction network and centrality methods. Some key genes asso-

收稿日期: 2022-05-10 修回日期: 2022-07-03 网络首发日期: 2022-09-07

*基金项目: 云南省创新引导与科技型企业培育计划 (202104BI090022); 云南省产业技术领军人才专项 (YNWR-CYJS-2018-056); 云南省重大科技专项计划项目 (202102AE090039)。

作者简介: 荣丽云 (1997—), 女, 湖南东安人, 在读硕士研究生, 主要从事动物分子遗传研究。

E-mail: 1554555085@qq.com

**通信作者 Corresponding authors: 陈强 (1980—), 男, 河南长垣人, 博士, 讲师, 主要从事动物分子遗传研究。E-mail: chq@sjtu.edu.cn; 鲁绍雄 (1972—), 男, 云南镇雄人, 博士, 教授, 主要从事动物遗传育种研究。E-mail: shxlu@263.net

网络首发地址: <https://kns.cnki.net/kcms/detail/53.1044.S.20220906.1718.017.html>



ciated with pork quality trait were identified by qPCR on three Saba and three Large White pigs.

[**Results**] Two gene co-expression modules including the blue and the red modules related to pork quality trait were identified, which the r values were 0.89 ($P=5E-6$) and 0.91 ($P=7E-7$), respectively. The genes in the blue module were mainly associated with transition between fast and slow fiber and extracellular matrix and so on. Among the identified genes, 10 genes were identified as the key genes associated with pork quality, including *COL3A1*, *COL5A1*, *COL1A2*, *COL14A1*, *CRTAP*, *COL5A3*, *LEPREL1*, *DEPTOR*, *COL4A2* and *COL6A2*. The genes in the red module were mainly related to skeletal muscle and fat cell differentiation. Among these genes, 11 genes were identified as the key genes associated with pork quality, including *FBXL4*, *RNF115*, *FBXO40*, *ASB4*, *ASB15*, *UBE2L6*, *RNF41*, *ASB5*, *HERC5*, *SOCS1* and *ASB11*. The expression levels of the top three genes in the blue module had extremely significant differences in longissimus dorsi muscles between Saba and Large white pig breeds ($P<0.01$), while the top three genes in the red module had no significant differences ($P>0.05$). [**Conclusion**] The results preliminarily identified the co-expression modules and key genes of meat quality between different types of pig breeds, but the key genes associated with meat quality trait and their effects between the pig breeds with the same meat type may be significantly different.

Keywords: pig; pork quality trait; gene expression; weighted gene co-expression network; key gene

肉质性状是养猪生产中最重要的经济性状之一,对养猪生产效益有着重要影响。提高猪肉品质一直是猪育种工作的重要目标,但由于肉质性状活体难以准确测定,且包括肌肉脂肪含量、大理石纹、嫩度、pH 值和滴水损失等诸多相互关联的子性状,常规育种方法改良难度大,难以获得理想的遗传进展。随着分子生物学及基因组学相关技术的发展,人们发掘了大量与猪肉品质性状相关的基因^[1-4]。尽管如此,关于猪肉品质性状的分子机制迄今仍不甚清楚。

基因共表达网络分析 (gene co-expression network analysis, GCNA) 是基于对同一性状的基因在表达模式上具有共表达特征这一设定,从构建的表达调控网络中发现并识别目标性状的相关基因。该方法对目标性状关键新基因的发掘及其功能预测具有重要作用。加权基因共表达网络分析 (weighted gene co-expression network analysis, WGCNA) 是目前开展 GCNA 最主要的方法^[5],人们利用该方法鉴定了大量与复杂性状相关的基因共表达模块和关键基因^[6-9]。

近年来,利用 GCNA 方法探索畜禽重要经济性状分子机制的研究逐渐增多,有关猪肉品质性状的基因共表达模块研究也有陆续报道^[10-13]。如 ZAPPATERRA 等^[2]对意大利大白猪半膜肌进行研究,识别了 4 个与肌肉脂肪含量相关的基因共表

达模块; ZHAO 等^[14]利用 WGCNA 法识别了与猪肉滴水损失相关的 *AASS*、*BCKDHB*、*ALDH6A1*、*MUT* 和 *MCCCI* 等 5 个基因。这些研究为进一步揭示猪肉品质性状形成的基因表达机制提供了有益探索,但目前关于肉质性状相关的基因共表达研究及可利用的基因共表达信息仍较少,尚需进一步深入研究。

相较于国外瘦肉型猪种,中国地方猪种肉质优良的遗传特性突出。由于两类猪种在肉质性状等方面存在较大差异,基于两类猪种的比较分析是揭示肉质性状分子机制的一种有效途径。本研究基于公共 GEO 数据库中大蒲莲猪 (中国地方品种) 和长白猪 (国外瘦肉型品种) 背最长肌基因表达数据集,通过 WGCNA、蛋白质与蛋白质相互作用 (protein-protein interaction, PPI) 网络及中心性分析等方法,挖掘与肉质性状相关的关键基因共表达模块和关键基因,并对识别的关键基因表达进行 qPCR 验证,以探索猪肉品质性状形成的分子机制,为联合运用经典育种和分子育种手段改良猪肉品质性状奠定基础并提供科学依据。

1 材料与方法

1.1 数据来源

数据来源于美国生物技术信息中心 (NCBI)

基因表达数据库(GEO, <https://www.ncbi.nlm.nih.gov/geo/profiles/>) 中的数据集 GSE119368, 包括 8 头大蒲莲猪(中国地方品种)和 8 头长白猪(国外瘦肉型品种) 6 月龄背最长肌基因表达数据。猪群饲料营养相同、饲养管理条件相近。基因表达数据由 Illumina HiSeq 2500 测序平台产生。

1.2 差异表达基因分析

差异表达基因分析利用基于 R 软件的 limma 包^[15](version 3.44.3) 进行, 以矫正的 $P < 0.05$ 和 $|\log_2 \text{fold change}| > 1$ 作为基因差异表达的显著性阈值。差异表达基因(differentially expressed genes, DEGs) 分布的火山图和表达热图采用 ggplot2 软件包(<https://cran.r-project.org/web/packages/ggplot2/>, version 3.3.3) 绘制。

1.3 基因共表达模块分析

基因共表达模块分析采用 WGCNA 方法^[5], 使用基于 R 环境的 WGCNA 软件包(<https://cran.r-project.org/web/packages/WGCNA/>, version 1.70-3) 进行。以大蒲莲猪为处理组、长白猪为对照组, 质控后获得的 12 494 个基因按如下流程进行分析:

(1) 计算两两基因间的皮尔逊相关系数(r), 并据此构建基因表达相似矩阵; (2) 将基因表达相似矩阵转换成邻接矩阵, 并通过软阈值 β ($\beta = 10$, 无尺度拓扑标准 $R^2 > 0.85$) 产生增强邻接矩阵; (3) 使用皮尔逊相关分析构建增强邻接矩阵的无监督基因共表达关系矩阵, 并基于基因间的关联程度指标拓扑重叠尺度(topological overlap measure, TOM) 将其转化为拓扑矩阵; (4) 基于基因间相异度($1 - \text{TOM}$), 使用平均连锁层次聚类法将具有相似基因表达模式的基因归为同一个基因模块, 采用动态剪切算法从系统聚类树中识别基因共模块, 并将相似度高于 75% 的基因模块进行合并。

1.4 基因功能分析

差异表达基因及关键模块中基因的功能采用基因本体论(gene ontology, GO) 富集及京都基因和基因组百科全书(Kyoto encyclopedia of genes and genomes, KEGG) 通路富集方法, 利用在线的 DAVID 软件^[16](<https://david.ncifcrf.gov/>, version 6.8) 进行分析, 通路显著性富集阈值为 $P < 0.05$ 。

1.5 PPI 网络构建

关键模块中编码蛋白基因间的互作关系采用

PPI 网络进行分析, 互作关系信息来源于在线数据库 STRING^[17](<http://string-db.org>, version 11.5), 筛选标准为互作得分大于 0.7。PPI 网络采用 Cytoscape 软件^[18](<https://cytoscape.org/>, version 3.8.0) 构建。

1.6 高度关联模块及关键基因识别

PPI 网络中与肉质性状高度关联模块的分析采用分子复合检测(molecular complex detection, MCODE) 算法, 并利用 Cytoscape 软件中的 MCODE 插件^[19](version 1.5.1) 进行, 相应参数设置为 Degree Cutoff=2、Max Depth=100、K-Core=2 以及 Node Score Cutoff=0.2。关键基因的识别采用 Subgraph、Eigenvector、Information、Closeness、Betweenness、LAC 和 Network 等 7 种中心性方法, 并利用 Cytoscape 软件中的 CytoNCA 插件^[20](version 2.1.6) 进行, 其中, 7 种方法得分最高的前 10 个基因即为关键基因。

1.7 关键基因验证

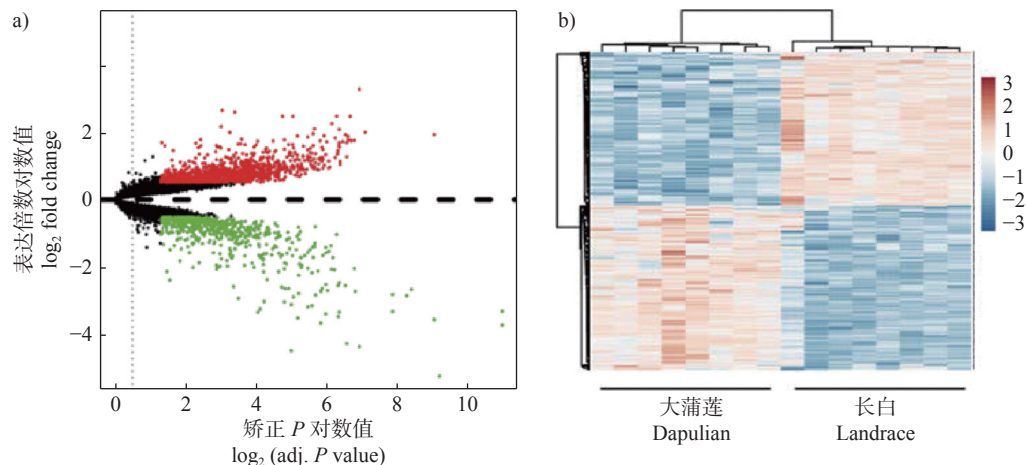
选择彼此无亲缘关系的地方品种撒坝猪和引进瘦肉型品种大白猪各 3 头, 在相同饲养管理条件下饲养至体质量为 90~100 kg 时屠宰, 采集背最长肌组织用于肉质关键基因验证。从每个识别的肉质性状高度关联模块中各选择 3 个关键基因, 采用 qPCR 方法检测关键基因在撒坝猪和大白猪背最长肌组织中的表达水平, 并采用非配对试验的 t 检验进行差异分析, 差异显著和极显著阈值分别为 $P < 0.05$ 和 $P < 0.01$ 。

2 结果与分析

2.1 DEGs 识别及功能分析

差异基因表达分析结果如图 1 所示。在大蒲莲猪和长白猪的背最长肌组织中共识别出 390 个 DEGs, 其中, 大蒲莲猪中显著高表达基因 189 个、显著低表达基因 201 个。

GO 分析结果(表 1) 显示: 189 个高表达基因极显著富集在 3 个 GO 生物学过程、3 个 GO 细胞组分和 2 个 GO 分子功能上, 201 个低表达基因极显著富集在 12 个 GO 生物学过程、5 个 GO 细胞组分和 2 个 GO 分子功能上。高表达基因主要与胶原纤维组织、成肌细胞融合和 PI3K-Akt 信号传导等生物学过程相关, 低表达基因主要与骨骼肌细胞分化、蛋白质折叠及神经元凋亡过程



注: a) 差异表达基因的分布, 红色和绿色点分别代表高表达和低表达基因; b) 差异表达基因的表达水平, 由红到蓝表示基因表达水平由高到低。
Note: a) distribution of differentially expressed genes, the red and the green dots indicate highly and lowly expressed genes, respectively; b) expression of differentially expressed genes, the color red to blue indicates high to low level of gene expression.

图 1 大蒲莲和长白猪差异表达基因分布及表达水平

Fig. 1 Distribution and expression level of differentially expressed genes in Dapulian and Landrace pigs

表 1 差异表达基因极显著富集的 GO 条目

Tab. 1 GO terms with extremely significant enrichment of differentially expressed genes

表达水平 expression level	GO 登录号 GO accession	条目名称 term description	富集基因数 gene number	条目类型 term type	P 值 P-value
高 high	GO: 0030199	胶原纤维组织 collagen fibril organization	5	BP	1.37E-4
	GO: 0007520	成肌细胞融合 myoblast fusion	3	BP	5.83E-3
	GO: 0014065	磷脂酰肌醇3-激酶信号 phosphatidylinositol 3-kinase signaling	3	BP	9.88E-3
	GO: 0005578	蛋白质细胞外基质 proteinaceous extracellular matrix	12	CC	4.14E-7
	GO: 0005604	基膜 basement membrane	4	CC	3.77E-3
	GO: 0005615	细胞外空间 extracellular space	17	CC	8.85E-3
	GO: 0050840	胞外基质结合 extracellular matrix binding	4	MF	1.25E-3
	GO: 0005509	钙离子结合 calcium ion binding	14	MF	4.77E-3
低 low	GO: 0035914	骨骼肌细胞分化 skeletal muscle cell differentiation	7	BP	1.29E-6
	GO: 0006457	蛋白折叠 protein folding	8	BP	2.69E-4
	GO: 0043524	神经元凋亡过程负调控 negative regulation of neuron apoptotic process	7	BP	3.77E-4
	GO: 0043065	凋亡过程的正调控 positive regulation of apoptotic process	7	BP	2.57E-3
	GO: 0006351	转录, DNA模板化 transcription, DNA-templated	12	BP	2.65E-3
	GO: 0090084	包涵体组装负调节 negative regulation of inclusion body assembly	3	BP	2.85E-3
	GO: 0009408	热反应 response to heat	4	BP	4.46E-3
	GO: 0051091	序列特异性DNA结合转录因子活性正调控 positive regulation of sequence-specific DNA binding transcription factor activity	5	BP	5.65E-3
	GO: 0043154	参与凋亡过程的半胱氨酸型内肽酶活性负调控 negative regulation of cysteine-type endopeptidase activity involved in apoptotic process	4	BP	7.64E-3
	GO: 0042692	肌肉细胞分化 muscle cell differentiation	3	BP	9.96E-3
	GO: 0097193	内在凋亡信号通路 intrinsic apoptotic signaling pathway	3	BP	9.96E-3
	GO: 0001894	组织内稳态 tissue homeostasis	3	BP	9.96E-3
	GO: 0005634	细胞核 nucleus	50	CC	1.41E-5
	GO: 0005829	细胞溶质 cytosol	25	CC	1.71E-5
	GO: 0005737	细胞质 cytoplasm	43	CC	3.01E-3
	GO: 0031674	I带 I band	3	CC	5.64E-3
	GO: 0031012	细胞外基质 extracellular matrix	6	CC	6.19E-3
	GO: 0019888	蛋白磷酸酶调节活性 protein phosphatase regulator activity	3	MF	2.38E-3
	GO: 0043565	序列特异性DNA结合 sequence-specific DNA binding	10	MF	8.63E-3

注: BP、CC和MF分别表示生物学过程、细胞组分和分子功能。

Note: BP, CC and MF indicates biological process, cellular component and molecular function, respectively.

调控等生物学过程相关。

KEGG 通路分析结果 (表 2) 显示: 189 个高表达基因和 201 个低表达基因分别极显著富集在 5 个和 3 个 KEGG 通路上。其中, 高表达基因主要与蛋白质消化吸收、醛固酮调节的钠重吸收以及心肌细胞的肾上腺素能信号等通路相关, 低表达基因主要与 MAPK 信号通路、内质网蛋白质加工以及雌激素信号通路等相关。

表 2 差异表达基因极显著富集的 KEGG 通路

Tab. 2 KEGG pathways with extremely significant enrichment of differentially expressed genes

表达水平 expression level	通路ID pathway ID	通路名称 pathway term	富集基因数 gene number	P 值 P-value
高 high	ssc04974	蛋白质消化吸收 protein digestion and absorption	6	1.02E-3
	ssc04960	醛固酮调节的钠重吸收 aldosterone-regulated sodium reabsorption	4	7.18E-3
	ssc04261	心肌细胞的肾上腺素能信号传导 adrenergic signaling in cardiomyocytes	6	8.37E-3
	ssc00910	氮代谢 nitrogen metabolism	3	9.00E-3
	ssc04068	FoxO 信号通路 FoxO signaling pathway	6	9.20E-3
低 low	ssc04010	MARK 信号通路 MAPK signaling pathway	10	2.67E-3
	ssc04141	内质网蛋白加工 protein processing in endoplasmic reticulum	8	3.06E-3
	ssc04915	雌激素信号通路 estrogen signaling pathway	6	6.03E-3

2.2 基因共表达模块分析

聚类结果显示: 大蒲莲猪和长白猪样本均各自聚成一类 (图 2a)。根据计算的基因间相异度, 识别了 6 个基因共表达模块 (图 2b); 合并相似性

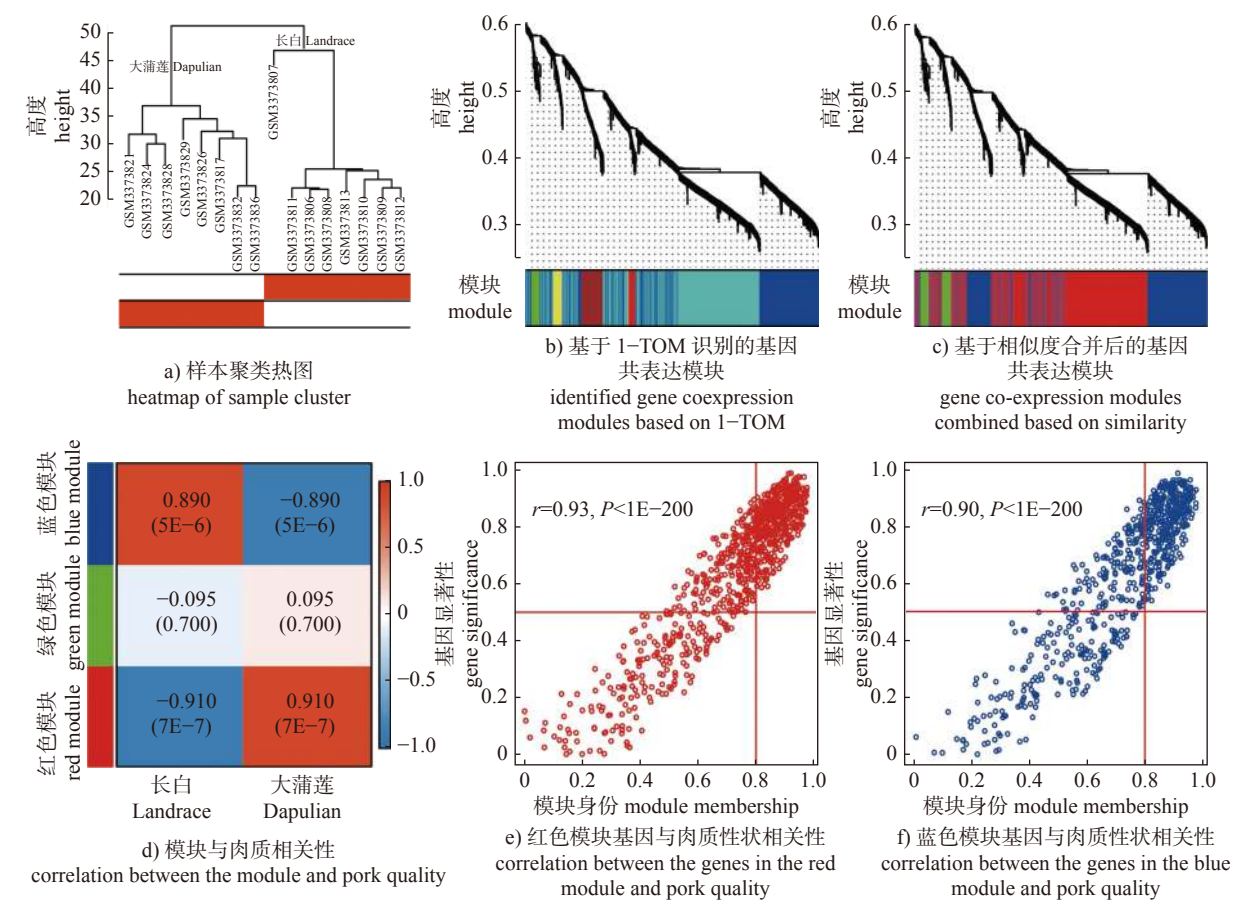


图 2 WGCNA 分析结果

Fig. 2 The results of WGCNA analysis

高于 75% 的模块后, 获得红色、蓝色和绿色 3 个基因共表达模块 (图 2c)。其中, 红色模块 ($r=0.91$, $P=7E-7$) 和蓝色模块 ($r=0.89$, $P=5E-6$) 与肉质性状的关联度最高 (图 2d), 两模块中基因表达与肉质均呈极显著相关 ($r=0.93$, $P<1E-200$; $r=0.90$, $P<1E-200$) (图 2e 和 2f)。

2.3 与肉质性状最显著关联的模块基因功能分析
与肉质性状最显著关联的红色模块 660 个基因极显著 ($P<0.01$) 富集到 17 个 GO 生物学过程和 12 个 KEGG 通路上 (显著性最高的前 5 个生物

学过程和 KEGG 通路见表 3), 主要与蛋白质折叠、骨骼肌细胞分化、糖原代谢过程和雌激素信号、蛋白质加工以及蛋白激酶等信号通路相关; 蓝色模块 527 个基因极显著 ($P<0.01$) 富集到 12 个 GO 生物学过程和 10 个 KEGG 通路上 (显著性最高的前 5 个生物学过程和 KEGG 通路见表 4), 主要与肌肉组织形态形成、心脏肌肉收缩、快慢纤维转换过程和蛋白质消化吸收、ECM-受体相互作用以及黏着斑等信号通路相关。

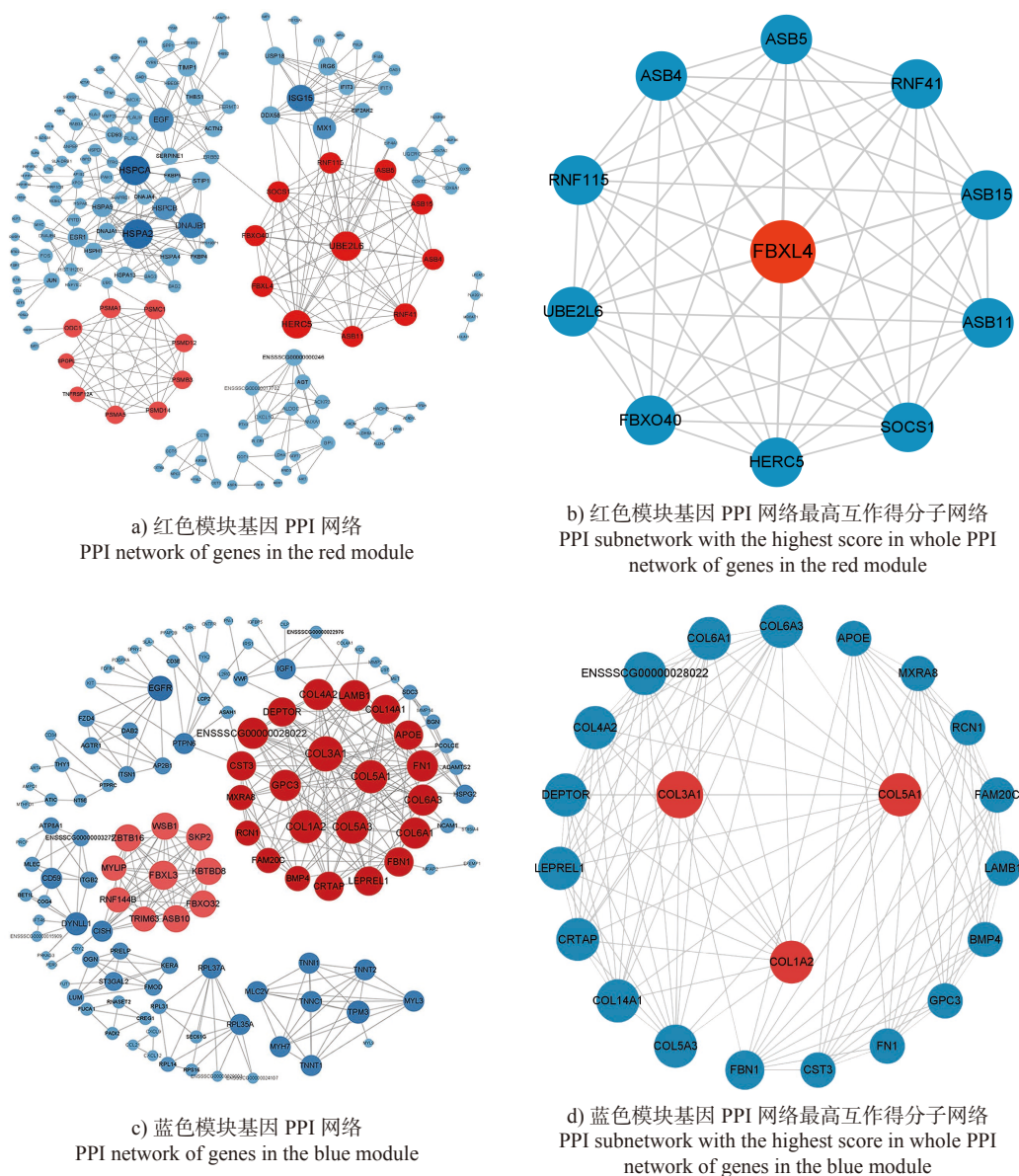
表 3 红色模块基因极显著富集的前 5 个 GO 条目和 KEGG 通路				
Tab. 3 Top 5 GO terms and KEGG pathways extremely significantly enriched by the genes in the red module				
富集方式 enriched method	GO 登录号/通路 ID GO accession/pathway ID	条目名称 term description	富集基因数 enriched gene number	P 值 P-value
GO 条目 GO term	GO: 0006457	蛋白质折叠 protein folding	19	2.72E-8
	GO: 0035914	骨骼肌细胞分化 skeletal muscle cell differentiation	8	7.19E-5
	GO: 0043086	催化活性负调控 negative regulation of catalytic activity	7	1.24E-4
	GO: 0005977	糖原代谢过程 glycogen metabolic process	6	2.70E-4
	GO: 0045444	脂肪细胞分化 fat cell differentiation	8	1.90E-3
KEGG 通路 KEGG pathway	ssc05164	甲型流感 influenza A	23	2.07E-6
	ssc01130	抗生素生物合成 biosynthesis of antibiotics	24	5.76E-6
	ssc04915	雌激素信号通路 estrogen signaling pathway	14	2.00E-4
	ssc04141	内质网蛋白加工 protein processing in endoplasmic reticulum	17	1.18E-3
	ssc04010	MARK 信号通路 MAPK signaling pathway	22	1.65E-3

表 4 蓝色模块基因极显著富集的前 5 个 GO 条目和 KEGG 通路				
Tab. 4 Top 5 GO terms and KEGG pathways extremely significantly enriched by the genes in the blue module				
富集方式 enriched method	GO 登录号/通路 ID GO accession/pathway ID	条目名称 term description	富集基因数 enriched gene number	P 值 P-value
GO 条目 GO term	GO: 0055010	心室心肌组织形态发生 ventricular cardiac muscle tissue morphogenesis	6	1.18E-4
	GO: 0060048	心肌收缩 cardiac muscle contraction	6	2.20E-4
	GO: 0014883	快慢肌纤维间转换 transition between fast and slow fiber	4	3.31E-4
	GO: 0043588	皮肤发育 skin development	6	4.82E-4
	GO: 0007160	细胞基质黏附 cell-matrix adhesion	8	6.75E-4
KEGG 通路 KEGG pathway	ssc04974	蛋白质消化吸收 protein digestion and absorption	14	9.75E-8
	ssc04512	ECM 受体互作 ECM-receptor interaction	13	8.71E-7
	ssc04510	黏着斑 focal adhesion	18	8.29E-6
	ssc04151	PI3K-Akt 信号通路 PI3K-Akt signaling pathway	22	8.59E-5
	ssc04260	心肌收缩 cardiac muscle contraction	9	5.27E-4

2.4 PPI 网络构建及关键基因识别
所构建的红色模块基因 PPI 网络包括 163 个基因 (节点) 和 371 条边 (图 3a), 最高得分的子网络包含 11 个基因 (*FBXL4*、*RNF115*、*FBXO40*、*ASB4*、*ASB15*、*UBE2L6*、*RNF41*、*ASB5*、*HERC5*、

SOCS1、*ASB11*) 和 55 条边 (图 3b), 11 个基因的中心性综合得分相同 (表 5), 为肉质性状相关的关键基因。

所构建的蓝色模块基因 PPI 网络包含 128 个基因和 341 条边 (图 3c), 最高得分的子网络包含



注: 红色和蓝色节点分别表示在长白猪中高表达和低表达基因, 节点越大代表连接数越多。

Note: Red and blue nodes represent the genes of high and low expression of Landrace pigs, respectively, which the bigger node represents more links.

图 3 基因互作网络与高关联度模块识别

Fig. 3 Gene interaction network and highly related module identification

22 个基因和 112 条边 (图 3d), 中心性综合得分最高的前 10 个基因为 *COL3A1*、*COL5A1*、*COL1A2*、*COL14A1*、*CRTAP*、*COL5A3*、*LEPREL1*、*DEPTOR*、*COL4A2* 和 *COL6A2* (表 6), 为肉质性状相关的关键基因。

2.5 关键基因表达分析及 qPCR 验证

由图 4 可知: 2 个模块各挑选的 3 个肉质性状关键基因在大蒲莲猪和长白猪背最长肌中的表达水平均呈极显著差异, 除 *RNF115* ($P=0.00184$) 外, *COL3A1* ($P=0.00016$)、*COL5A1* ($P=0.00165$)、

COL1A2 ($P=0.000073$)、*FBXL4* ($P=0.00097$) 和 *FBXO40* ($P=0.00054$) 5 个基因的表达水平均呈现为长白猪高于大蒲莲猪。

qPCR 结果 (图 5) 显示: 蓝色模块中的 *COL3A1* ($P<0.0001$)、*COL5A1* ($P=0.0006$) 和 *COL1A2* ($P<0.0001$) 在撒坝猪和大白猪背最长肌中的表达水平呈极显著差异, 而红色模块中 3 个基因 (*RNF115*、*FBXL4* 和 *FBXO40*) 的表达水平在 2 个品种间则无显著差异 ($P>0.05$)。

表 5 红色模块基因 PPI 子网络基因中心性分析

Tab. 5 Centrality analysis of genes in the PPI sub-network of the red module

序号 rank	基因 gene	子图 subgraph	特征向量 eigenvector	信息 information	本地平均连通性 local average connectivity	介性 betweenness	近性 closeness	网络 network
1	<i>FBXL4</i>	2002.74	0.301 5	6.049 9	9	0	1	10
2	<i>RNF115</i>	2002.74	0.301 5	6.049 9	9	0	1	10
3	<i>FBXO40</i>	2002.74	0.301 5	6.049 9	9	0	1	10
4	<i>ASB4</i>	2002.74	0.301 5	6.049 9	9	0	1	10
5	<i>ASB15</i>	2002.74	0.301 5	6.049 9	9	0	1	10
6	<i>UBE2L6</i>	2002.74	0.301 5	6.049 9	9	0	1	10
7	<i>RNF41</i>	2002.74	0.301 5	6.049 9	9	0	1	10
8	<i>ASB5</i>	2002.74	0.301 5	6.049 9	9	0	1	10
9	<i>HERC5</i>	2002.74	0.301 5	6.049 9	9	0	1	10
10	<i>SOCS1</i>	2002.74	0.301 5	6.049 9	9	0	1	10
11	<i>ASB11</i>	2002.74	0.301 5	6.049 9	9	0	1	10

表 6 蓝色模块基因 PPI 子网络基因中心性分析

Tab. 6 Centrality analysis of genes in the PPI sub-network of the blue module

序号 rank	基因 gene	子图 subgraph	特征向量 eigenvector	信息 information	本地平均连通性 local average connectivity	介性 betweenness	近性 closeness	网络 network
1	<i>COL3A1</i>	5196.21	0.292 90	6.244 7	9.17	220	0.700	11
2	<i>COL5A1</i>	5015.27	0.287 85	6.025 4	10.00	0	0.525	11
3	<i>COL1A2</i>	5015.27	0.287 85	6.025 4	10.00	0	0.525	11
4	<i>COL14A1</i>	5015.27	0.287 85	6.025 4	10.00	0	0.525	11
5	<i>CRTAP</i>	5015.27	0.287 85	6.025 4	10.00	0	0.525	11
6	<i>COL5A3</i>	5015.27	0.287 85	6.025 4	10.00	0	0.525	11
7	<i>LEPREL1</i>	5015.27	0.287 85	6.025 4	10.00	0	0.525	11
8	<i>DEPTOR</i>	5015.27	0.287 85	6.025 4	10.00	0	0.525	11
9	<i>COL4A2</i>	5015.27	0.287 85	6.025 4	10.00	0	0.525	11
10	<i>COL6A2</i>	5015.27	0.287 85	6.025 4	10.00	0	0.525	11
11	<i>COL6A3</i>	5015.26	0.287 85	6.025 4	10.00	0	0.525	11
12	<i>COL6A1</i>	5015.26	0.287 85	6.025 4	10.00	0	0.525	11
13	<i>FBN1</i>	907.06	0.038 20	5.785 3	7.20	216	0.656	9
14	<i>FN1</i>	819.99	0.011 81	5.521 3	8.00	0	0.477	9
15	<i>CST3</i>	819.99	0.011 81	5.521 3	8.00	0	0.477	9
16	<i>APOE</i>	819.99	0.011 81	5.521 3	8.00	0	0.477	9
17	<i>MXRA8</i>	819.99	0.011 81	5.521 3	8.00	0	0.477	9
18	<i>RCN1</i>	819.99	0.011 81	5.521 3	8.00	0	0.477	9
19	<i>FAM20C</i>	819.99	0.011 81	5.521 3	8.00	0	0.477	9
20	<i>LAMB1</i>	819.99	0.011 81	5.521 3	8.00	0	0.477	9
21	<i>BMP4</i>	819.99	0.011 81	5.521 3	8.00	0	0.477	9
22	<i>GPC3</i>	819.99	0.011 81	5.521 3	8.00	0	0.477	9

3 讨论

本研究基于 WGCNA 方法识别了 2 个与猪肉质性状显著关联的基因共表达模块，并通过构建的模块基因间互作关系网络分别识别出 10 个和

11 个关键基因。

在蓝色模块识别的 10 个猪肉质性状关键基因中，有 7 个基因 (*COL3A1*、*COL5A1*、*COL1A2*、*COL14A1*、*COL5A3*、*COL4A2* 和 *COL6A2*) 为胶原蛋白家族相关基因，该家族基因编码一类胶原

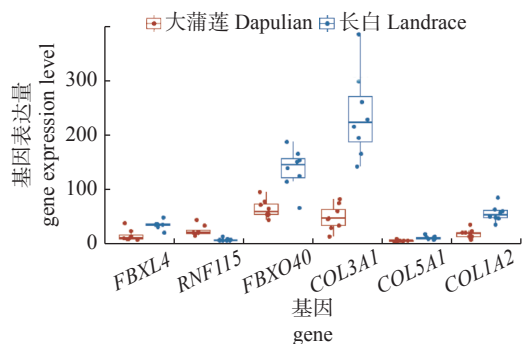


图4 大蒲莲和长白猪背最长肌中6个关键基因的表达水平

Fig. 4 The expression levels of six key genes in longissimus dorsi muscle of Dapulian and Landrace pigs

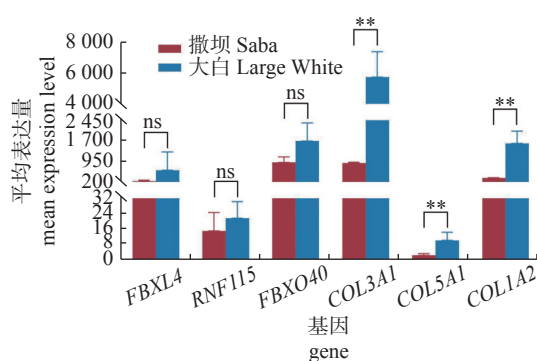


图5 6个关键基因在撒坝和大白猪背最长肌中的表达水平

Fig. 5 The expression of six genes in longissimus dorsi muscle of Saba and Large White pigs

蛋白,主要存在于动物结缔组织中。有研究表明:胶原蛋白对动物肌肉脂肪形成、肌肉嫩度、系水力及滴水损失等具有重要影响^[21-24]。NAKAJIMA等^[25]研究了牛胸最长肌的肌肉脂肪,发现胶原蛋白V和VI型对肌肉脂肪的形成具有正向调控作用;曾勇庆等^[26]对莱芜猪背最长肌的胶原蛋白含量及性质进行分析后发现:随着胶原蛋白含量的逐渐增加,猪肉的大理石纹和持水性能逐渐增加。

已有研究^[27-29]表明:COL3A1、COL5A1、COL1A2和COL14A1等4个胶原蛋白家族基因与猪肉品质性状相关。LIM等^[27]对巴克夏猪背最长肌转录组的研究显示:COL1A2、COL5A1和COL14A1与猪肌肉肌肉脂肪含量显著相关;BAO等^[28]对大白猪杂交群体的研究发现:COL3A1基因的表达上调有利于猪肌纤维的生成;FERNÁNDEZ-

EZ-BARROSO等^[29]对伊比利亚猪的研究发现:COL14A1基因对猪肉嫩度有显著影响。可见,胶原蛋白家族基因可能主要与猪肉的肌肉脂肪含量和嫩度等相关。本研究的qPCR验证结果表明:COL3A1、COL5A1和COL1A2基因在肉质差异较大的地方品种撒坝猪和引进品种大白猪背最长肌中的表达水平差异极显著,进一步提示上述基因对肉质性状有着潜在重要影响。而有关COL5A3、COL4A2和COL6A2基因对猪肉品质性状的影响迄今未见相关研究报道,其确切功能尚需进一步研究。此外,本研究识别的另外3个关键基因(CRTAP、LEPREL1和DEPTOR)中,CRTAP是软骨关联蛋白编码基因,主要与骨质发育相关,有研究表明该基因参与胶原形成和胞外基质降解^[30],而LEPREL1和DEPTOR基因对肉质性状的影响目前仍未见报道,2个基因与肉质的确切相关性仍需进一步通过功能试验确认。

在红色模块识别的11个关键基因中,已有研究^[31-37]显示其中的一些基因与肉质性状相关,如FBXL4、FBXO40、ASB4和ASB15等。FBXO40为F-box蛋白家族成员之一,该基因的敲除可显著促进猪骨骼肌纤维增粗^[31-32]。FBXL4是一个与线粒体功能控制有关的核基因,与FBXO40同属于F-box蛋白家族,当FBXL4基因敲除时小鼠出现线粒体功能障碍及体质量减轻等现象^[33]。ASB4、ASB15和ASB5同属ASB基因家族成员,在骨骼肌和心肌中呈高表达,被认为是猪肉品质性状遗传改良的潜在候选基因^[34-35]。UBE2L6在动物脂肪细胞中可作为介导脂肪分解的负调节因子促进高脂饮食诱导产生肥胖,敲除该基因能够抑制脂肪生成和减少脂肪细胞数量^[36-37]。目前,关于RNF41、RNF115、HERC5和SOCS1基因与肉质性状的相关性尚未见报道。

尽管撒坝猪和大白猪背最长肌中FBXL4和FBXO40基因的表达差异与大蒲莲和长白猪的趋势一致,但RNF115基因的表达差异则呈相反趋势,且这3个关键基因在撒坝猪和大白猪中的表达水平差异均不显著。究其原因,可能在于撒坝猪和大蒲莲猪虽同为肉质优良的中国地方猪种,长白猪和大白猪均为国外瘦肉型猪种,但同类猪种在部分肉质指标上仍存较大差异,如体质量为100 kg的撒坝猪肌肉脂肪含量和肉色评分分别为

(6.53±1.65)% 和 (3.19±0.31)分^[38], 而大蒲莲猪则分别为 (4.63±0.88)% 和 (4.25±0.14)分^[39], 可见其肉质形成的分子机制可能存在一定差异, 这可能也是造成验证结果不一致的原因之一。如能在不同猪种中进一步开展表达水平验证, 重点比较不同地方品种(如撒坝猪和大蒲莲猪)的基因表达差异情况, 深入揭示这些基因影响的具体性状及其效应, 将有助于猪肉品质性状的基因发掘并为精准育种提供更有价值的分子标记。

4 结论

本研究基于 WGCNA 方法识别了与猪肉品质性状的 2 个关键基因共表达模块和 21 个关键基因, 其中 7 个胶原蛋白基因家族相关基因 *COL3A1*、*COL5A1*、*COL1A2*、*COL14A1*、*COL5A3*、*COL4A2* 和 *COL6A2* 在猪肉品质性状形成中可能有重要作用。大蒲莲猪和撒坝猪虽同为地方猪种, 但其肉质性状形成的分子机制可能存在较大差异。

[参考文献]

- [1] DAMON M, WYSZYNSKA-KOKO J, VINCENT A, et al. Comparison of muscle transcriptome between pigs with divergent meat quality phenotypes identifies genes related to muscle metabolism and structure[J]. PLoS One, 2012, 7(3): e33763. DOI: [10.1371/journal.pone.0033763](https://doi.org/10.1371/journal.pone.0033763).
- [2] ZAPPATERA M, GIOIOSA S, CHILLEMI G, et al. Muscle transcriptome analysis identifies genes involved in ciliogenesis and the molecular cascade associated with intramuscular fat content in Large White heavy pigs[J]. PLoS One, 2020, 15(5): e0233372. DOI: [10.1371/journal.pone.0233372](https://doi.org/10.1371/journal.pone.0233372).
- [3] YU K F, SHU G, YUAN F F, et al. Fatty acid and transcriptome profiling of longissimus dorsi muscles between pig breeds differing in meat quality[J]. International Journal of Biological Sciences, 2013, 9(1): 108. DOI: [10.7150/ijbs.5306](https://doi.org/10.7150/ijbs.5306).
- [4] YAO C G, PANG D X, LU C, et al. Data mining and validation of AMPK pathway as a novel candidate role affecting intramuscular fat content in pigs[J]. Animals, 2019, 9(4): 137. DOI: [10.3390/ani9040137](https://doi.org/10.3390/ani9040137).
- [5] LANGFELDER P, HORVATH S. WGCNA: an R package for weighted correlation network analysis[J]. BMC Bioinformatics, 2008, 9(1): 559. DOI: [10.1186/1471-2105-9-559](https://doi.org/10.1186/1471-2105-9-559).
- [6] XING K, LIU H T, ZHANG F X, et al. Identification of key genes affecting porcine fat deposition based on co-expression network analysis of weighted genes[J]. Journal of Animal Science and Biotechnology, 2021, 12: 100. DOI: [10.1186/s40104-021-00616-9](https://doi.org/10.1186/s40104-021-00616-9).
- [7] GONG Y, HE J, LI B, et al. Integrated analysis of lncRNA and mRNA in subcutaneous adipose tissue of Ningxiang pig[J]. Biology, 2021, 10(8): 726. DOI: [10.3390/biology10080726](https://doi.org/10.3390/biology10080726).
- [8] KOGELMAN L, CIRERA S, ZHERNAKOVA D V, et al. Identification of co-expression gene networks, regulatory genes and pathways for obesity based on adipose tissue RNA Sequencing in a porcine model[J]. BMC Medical Genomics, 2014, 7: 57. DOI: [10.1186/1755-8794-7-57](https://doi.org/10.1186/1755-8794-7-57).
- [9] HE Y Q, MALTECCA C, TIEZZI F, et al. Transcriptome analysis identifies genes and co-expression networks underlying heat tolerance in pigs[J]. BMC Genetics, 2020, 21(1): 44. DOI: [10.1186/s12863-020-00852-4](https://doi.org/10.1186/s12863-020-00852-4).
- [10] PONSUKSILI S, DU Y, HADLICH F, et al. Correlated mRNAs and miRNAs from co-expression and regulatory networks affect porcine muscle and finally meat properties[J]. BMC Genomics, 2013, 14: 533. DOI: [10.1186/1471-2164-14-533](https://doi.org/10.1186/1471-2164-14-533).
- [11] PONSUKSILI S, SIENGDEE P, DU Y, et al. Identification of common regulators of genes in co-expression networks affecting muscle and meat properties[J]. PLoS One, 2015, 10(4): e0123678. DOI: [10.1371/journal.pone.0123678](https://doi.org/10.1371/journal.pone.0123678).
- [12] CHOTIMA P, CHRISTELLE H A, KANJANA T, et al. Revealing pathways associated with feed efficiency and meat quality traits in slow-growing chickens[J]. Animals, 2021, 11(10): 2977. DOI: [10.3390/ani11102977](https://doi.org/10.3390/ani11102977).
- [13] DINIZ W, MAZZONI G, COUTINHO L L, et al. Detection of co-expressed pathway modules associated with mineral concentration and meat quality in Nelore cattle[J]. Frontiers in Genetics, 2019, 10: 210. DOI: [10.3389/fgene.2019.00210](https://doi.org/10.3389/fgene.2019.00210).
- [14] ZHAO X, WANG C, WANG Y, et al. Weighted gene co-expression network analysis reveals potential candidate genes affecting drip loss in pork[J]. Animal Genetics, 2020, 51(6): 855. DOI: [10.1111/age.13006](https://doi.org/10.1111/age.13006).
- [15] RITCHIE M E, BELINDA P, WU D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies[J]. Nucleic Acids Research, 2015, 43(7): e47. DOI: [10.1093/nar/gkv007](https://doi.org/10.1093/nar/gkv007).
- [16] WEI H D, SHERMAN B T, LEMPICKI R A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists[J]. Nucleic Acids Research, 2009, 37(1): 1. DOI: [10.1093/nar/gkn923](https://doi.org/10.1093/nar/gkn923).
- [17] DAMIAN S, ANDREA F, STEFAN W, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life[J]. Nucleic Acids Research, 2015, 43: 447. DOI: [10.1093/nar/gku1003](https://doi.org/10.1093/nar/gku1003).
- [18] PAUL S, ANDREW M, OWEN O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks[J]. Genome Research, 2003, 13(11): 2498. DOI: [10.1101/gr.1239303](https://doi.org/10.1101/gr.1239303).
- [19] BADER G D, HOGUE C W. An automated method for finding molecular complexes in large protein interaction networks[J]. BMC Bioinformatics, 2003, 4: 2. DOI: [10.1186/1471-2105-4-2](https://doi.org/10.1186/1471-2105-4-2).

- 1186/1471-2105-4-2.
- [20] TANG Y, LI M, WANG J X, et al. CytoNCA: a cytoscape plugin for centrality analysis and evaluation of protein interaction networks[J]. *Biosystems*, 2015, 127: 67. DOI: [10.1016/j.biosystems.2014.11.005](https://doi.org/10.1016/j.biosystems.2014.11.005).
- [21] HAUSMAN G J. Meat science and muscle biology symposium: the influence of extracellular matrix on intramuscular and extramuscular adipogenesis[J]. *Journal of Animal Science*, 2012, 90(3): 942. DOI: [10.2527/jas.2011-4616](https://doi.org/10.2527/jas.2011-4616).
- [22] LUCARINI M, DURAZZO A, SCIUBBA F, et al. Stability of the meat protein type I collagen: influence of pH, ionic strength, and phenolic antioxidant[J]. *Foods*, 2020, 9(4): 480. DOI: [10.3390/foods9040480](https://doi.org/10.3390/foods9040480).
- [23] LEPETIT J. A theoretical approach of the relationships between collagen content, collagen cross-links and meat tenderness[J]. *Meat Science*, 2007, 76(1): 147. DOI: [10.1016/j.meatsci.2006.10.027](https://doi.org/10.1016/j.meatsci.2006.10.027).
- [24] MCBRYAN J, HAMILL R M, DAVEY G, et al. Identification of suitable reference genes for gene expression analysis of pork meat quality and analysis of candidate genes associated with the trait drip loss[J]. *Meat Science*, 2010, 86(2): 436. DOI: [10.1016/j.meatsci.2010.05.030](https://doi.org/10.1016/j.meatsci.2010.05.030).
- [25] NAKAJIMA I, MUROYA S, TANABE R, et al. Positive effect of collagen V and VI on triglyceride accumulation during differentiation in cultures of bovine intramuscular adipocytes[J]. *Differentiation*, 2002, 70(2/3): 84. DOI: [10.1046/j.1432-0436.2002.700203.x](https://doi.org/10.1046/j.1432-0436.2002.700203.x).
- [26] 曾勇庆, 王根林, 魏述东, 等. 莱茵猪肌肉胶原蛋白的发育性变化及其与肉质的相关性[J]. *中国农业科学*, 2008, 41(2): 619. DOI: [10.3864/j.issn.0578-1752.2008.02.043](https://doi.org/10.3864/j.issn.0578-1752.2008.02.043).
- [27] LIM K S, LEE K T, PARK J E, et al. Identification of differentially expressed genes in longissimus muscle of pigs with high and low intramuscular fat content using RNA sequencing[J]. *Animal Genetics*, 2017, 48(2): 166. DOI: [10.1111/age.12518](https://doi.org/10.1111/age.12518).
- [28] BAO X J, ZENG Y Q, WEI S D, et al. Developmental changes of Col3a1 mRNA expression in muscle and their association with intramuscular collagen in pigs[J]. *Journal of Genetics and Genomics*, 2007, 34(3): 223. DOI: [10.1016/S1673-8527\(07\)60023-X](https://doi.org/10.1016/S1673-8527(07)60023-X).
- [29] FERNÁNDEZ-BARROSO M, CARABALLO C, SILIÓ L, et al. Differences in the loin tenderness of Iberian pigs explained through dissimilarities in their transcriptome expression profile[J]. *Animals*, 2020, 10(9): 1715. DOI: [10.3390/ani10091715](https://doi.org/10.3390/ani10091715).
- [30] VALLI M, BARNES A M, GALLANTI A, et al. Deficiency of CRTAP in non-lethal recessive osteogenesis imperfecta reduces collagen deposition into matrix[J]. *Clinical Genetics*, 2011, 82(5): 453. DOI: [10.1111/j.1399-0004.2011.01794.x](https://doi.org/10.1111/j.1399-0004.2011.01794.x).
- [31] ZOU Y L, LI Z Y, ZOU Y J, et al. An *FBXO40* knockout generated by CRISPR/Cas9 causes muscle hypertrophy in pigs without detectable pathological effects[J]. *Biochemical and Biophysical Research Communications*, 2018, 498(4): 940. DOI: [10.1016/j.bbrc.2018.03.085](https://doi.org/10.1016/j.bbrc.2018.03.085).
- [32] SHI J, LUO L Q, EASH J, et al. The SCF-Fbxo40 complex induces IRS1 ubiquitination in skeletal muscle, limiting IGF1 signaling[J]. *Developmental Cell*, 2011, 21(5): 835. DOI: [10.1016/j.devcel.2011.09.011](https://doi.org/10.1016/j.devcel.2011.09.011).
- [33] ALSINA D, LYTOVCHENKO O, SCHAB A, et al. FBXL4 deficiency increases mitochondrial removal by autophagy[J]. *EMBO Molecular Medicine*, 2020, 12(7): e11659. DOI: [10.15252/emmm.201911659](https://doi.org/10.15252/emmm.201911659).
- [34] EHRLICH K C, LACEY M, EHRLICH M. Epigenetics of skeletal muscle-associated genes in the *ASB*, *LRRC*, *TMEM*, and *OSBPL* gene families[J]. *Epigenomes*, 2020, 4(1): 1. DOI: [10.3390/epigenomes4010001](https://doi.org/10.3390/epigenomes4010001).
- [35] VERBRUGGE S A J, SCHÖNFELDER M, BECKER L, et al. Genes whose gain or loss-of-function increases skeletal muscle mass in mice: a systematic literature review[J]. *Frontiers in Physiology*, 2018, 9: 553. DOI: [10.3389/fphys.2018.00553](https://doi.org/10.3389/fphys.2018.00553).
- [36] WEI W P, LI Y Q, LI Y Y, et al. Adipose-specific knockout of ubiquitin-conjugating enzyme E2L6 (Ube2l6) reduces diet-induced obesity, insulin resistance, and hepatic steatosis[J]. *Journal of Pharmacological Sciences*, 2021, 145(4): 327. DOI: [10.1016/j.jphs.2020.12.008](https://doi.org/10.1016/j.jphs.2020.12.008).
- [37] MARCELIN G, LIU S M, SCHWARTZ G J, et al. Identification of a loss-of-function mutation in Ube2l6 associated with obesity resistance[J]. *Diabetes*, 2013, 62(8): 2784. DOI: [10.2337/db12-1054](https://doi.org/10.2337/db12-1054).
- [38] 连林生, 鲁绍雄, 严达伟, 等. 撒坝猪配套系选育初报[J]. *猪业科学*, 2006, 6: 58. DOI: [10.3969/j.issn.1673-5358.2006.06.018](https://doi.org/10.3969/j.issn.1673-5358.2006.06.018).
- [39] 郭建凤, 吕山东, 王继英, 等. 大蒲莲猪与长蒲二元猪生长育肥性能、胴体性能及肉品质比较[J]. *畜牧与兽医*, 2010, 42(11): 24.

责任编辑: 何馨成