

网络出版时间:2022-04-20 10:15 DOI:10.13207/j.cnki.jnwafu.2022.10.001
网络出版地址:https://kns.cnki.net/kcms/detail/61.1390.S.20220418.1003.002.html

3 个肉牛群体新转录本注释及功能解析

王正文¹,张凌云¹,安雪姣¹,张小强²,杨雅楠¹,蔡原¹,赵生国¹

(1 甘肃农业大学 动物科学技术学院,甘肃 兰州 730070;2. 泾川县畜牧兽医中心,甘肃 平凉 744399)

【摘要】【目的】挖掘肉牛(*Bos taurus*)新转录本,探讨其潜在功能,为进一步探明牛肉品质差异的遗传机制提供思路。【方法】以平凉红牛、杂交和牛(和牛×平凉红牛)和西门塔尔牛为研究对象,采集其背最长肌,进行 RNA-seq 测序,构建 cDNA 文库,将重新组装的测序数据与牛参考基因组序列进行比对分析,优化已注释转录本的信息,挖掘新转录本并对其进行功能富集分析。随机挑选 9 个显著差异表达的新转录本进行 RT-qPCR 验证。【结果】对 1 081 个转录本的 5'非翻译区域(UTR)和 1 586 个转录本的 3'UTR 分别实现了延伸,并对 1 038 个转录本的 5'和 3'UTR 都实现了延伸。挖掘新转录本 442 个,其中 92 个新转录本在不同群体间表达差异显著($P < 0.05$);GO 分析显示,表达差异显著的新转录本在代谢过程、酶催化活性和细胞部分等显著富集;KEGG 富集发现,显著差异表达的新转录本显著富集在味觉转导、过氧化物酶、河马(Hippo)和 MAPK 等信号通路中。RT-qPCR 结果与 RNA-seq 结果在定量方面存在差异,但变化趋势一致,说明转录组测序结果可靠。【结论】新转录本通过各种代谢途径和通路广泛参与调控牛肌肉生长、脂肪沉积、疾病和免疫等各项生命活动。

【关键词】 平凉红牛;杂交和牛;西门塔尔牛;新转录本;肉质性状

【中图分类号】 S823.9⁺2

【文献标志码】 A

【文章编号】 1671-9387(2022)10-0001-13

Annotation and function analysis on new transcripts of three beef cattle groups

WANG Zhengwen¹, ZHANG Lingyun¹, AN Xuejiao¹, ZHANG Xiaoqiang²,
YANG Yanan¹, CAI Yuan¹, ZHAO Shengguo¹

(1 College of Animal Science and Technology, Gansu Agricultural University, Lanzhou, Gansu 730070, China;

2 Animal Husbandry and Veterinary Center of Jingchuan County, Pingliang, Gansu 744399, China)

Abstract: 【Objective】 This study discovered new transcripts of beef cattle (*Bos taurus*) and explored their potential functions to provide ideas for further research on genetic mechanism of beef quality differences. 【Method】 The longest dorsal muscle was collected for RNA-seq sequencing using Pingliang red cattle, crossbred Wagyu cattle (Wagyu×Pingliang Red Bull) and Simmental cattle as research subjects. The cDNA libraries were constructed, the reassembled sequencing data were compared with the bovine reference genome, the information of annotated transcripts was optimized, and new transcripts were searched to conduct functional enrichment analysis. Nine new transcripts with significant differences were randomly selected for RT-qPCR validation. 【Result】 Extensions were achieved for the 5' untranslated region (UTR) of 1 081 transcripts and the 3'UTR region of 1 586 transcripts, including both regions of 1 038 transcripts. A total of 442 new transcripts were mined, 92 of which were significantly ($P < 0.05$) different expressed a-

【收稿日期】 2021-10-25

【基金项目】 国家科技支撑计划项目(2015BAD03B04-4);甘肃省农业科技创新项目(GNCX-2014-32)

【作者简介】 王正文(1997-),男,甘肃临夏人,在读硕士,主要从事动物遗传育种研究。E-mail:gsauwangzhengwen@163.com

【通信作者】 赵生国(1976-),男,甘肃庆阳人,教授,博士,博士生导师,主要从事动物遗传学理论与应用研究。

E-mail:zhaosg@gsau.edu.cn

mong groups. GO analysis showed that new transcripts with significantly differential expression were significantly enriched in metabolic processes, enzyme catalytic activity and cell parts. KEGG enrichment revealed that these new transcripts were significantly enriched in taste transduction, peroxisome, Hippo signaling pathway, and MAPK signaling pathway. The RT-qPCR results differed from the RNA-seq results in terms of quantification, but the trends were consistent, indicating that the transcriptome sequencing results were reliable. 【Conclusion】 These new transcripts widely involved in regulating cattle muscle growth, fat deposition, disease, immunity and other life activities through various metabolic pathways and pathways.

Key words: Pingliang red cattle; crossbred Wagyu; Simmental cattle; new transcript; meat traits

为改良我国黄牛肉用性能, 研究人员先后引入了西门塔尔牛、和牛和南德温牛等国外优良肉牛品种(冻精)与本地牛进行杂交。西门塔尔牛具有体型大、生长快、肌肉脂肪含量低的特点^[1]。日本和牛肌肉脂肪含量丰富, 不饱和脂肪酸含量高, 肉质好^[2]。平凉红牛是以早胜牛为基础母牛, 经南德温牛等品种改良形成的一个肉用群体^[3], 其眼肌在大理石花纹、脂肪沉积及风味口感等方面与日本和牛接近^[4]。相关研究发现, 和牛与平凉红牛杂交(和牛♂×平凉红牛♀)后代的肉品质具有一定的优势^[5]。肌肉脂肪沉积与脂肪酸组成等肉品质特性受品种和营养等因素的影响, 这些因素对肉品质的影响与机体中相关基因及信号通路有关^[6]。近年来, 人们采用高通量测序技术获得了基因组、表观基因组、转录组、蛋白质组和代谢组等不同来源的生物信息学数据, 这些数据正引领人们进入“大数据”时代^[7]。生物信息学分析作为大量生物信息学数据与生物表型和功能机制的桥梁而备受青睐^[8], 许多生物信息学工具应用于单核苷酸多态性研究^[9]、差异表达基因筛选^[10]及新转录本挖掘和基因功能注释等方面^[11-12]。转录组测序和生物信息学的发展, 使得与肉品质相关的候选基因与通路不断被挖掘出来。Xing 等^[13]利用加权基因共表达网络分析(WGCNA)鉴定出 *RAD9A*、*TCAP*、*GPS2* 和 *APOF* 是调控猪(*Sus scrofa*)脂肪沉积的关键基因。Zhang 等^[14]研究发现, miR-22-3p 靶向调节基因 *WFIKKN2* 参与了牛脂肪细胞的分化。Chen 等^[15]研究表明, 抑制 Wnt/ β -catenin 信号通路可以促进猪前体肌肉脂肪的分化并改善脂肪酸组成。Rizzo 等^[16]研究发现, AMPK (AMP-activated protein kinase) 是调控肌肉脂肪合成的一条重要的信号通路。目前, RNA-seq 技术在新转录本挖掘和基因结构优化方面受到关注, 前人采用该技术已经在蜜蜂^[17]、羊布鲁氏菌^[18]和虹鳟^[19]上挖掘出新基因, 并对已知编码区结构进行了延伸优化, 这说明现有数据库中对转录本的注释还

不全面。为了完善牛肉品质差异的基因组注释信息, 本研究以平凉红牛、杂交和牛和西门塔尔牛为对象, 通过 RNA-seq 技术对新转录本进行了预测, 并对其进行了注释和功能解析, 以探索新转录本调控牛肌肉生长和脂肪沉积等各项生命活动的途径, 为探明肉牛肉品质差异的遗传机制提供参考依据。

1 材料与方法

1.1 试验动物

在甘肃省泾川旭康食品有限责任公司选取 6 月龄左右、健康状况良好、体质量为 (160.24 ± 5.36) kg/头的平凉红牛(P 组)、杂交和牛(和牛♂×平凉红牛♀, J 组)和西门塔尔牛(S 组)各 24 头, 每组 3 个重复, 每个重复 8 头牛。3 组试验牛全部采用全混合日粮(TMR)模式饲喂, 于 08:00, 18:00 各饲喂 1 次, 自由饮水。根据 NRC(2003)肉牛营养标准, 设计 4 个阶段(体质量 150~250, 250~350, 350~450, 450~550 kg)的基础日粮(表 1)。试验牛统一饲养 540 d 后, 每组随机选取 6 头体质量相近((587.8 ± 65.36) kg/头)的牛进行屠宰, 屠宰后, 取背最长肌样本, 装入冻存管后置入液氮, 带回实验室, 于 -80 °C 保存备用。

1.2 牛 RNA 提取、cDNA 文库构建及测序

使用 TRIzol 试剂盒(Invitrogen, Carlsbad, CA, USA)从每个样品中提取总 RNA, 用 Nano-Drop2000 (Thermo Scientific, Waltham, MA, USA)和 Agilent 2100 (Agilent, Santa Clara, CA, USA)检测 RNA 样品的浓度和质量, 将同组 6 个样本的 RNA 任意 2 个等质量混合为一个样本后, 用 Oligo(dT)磁珠(百迈客生物科技有限公司, 无锡)富集 mRNA, 利用片段缓冲液将 mRNA 切成短片段, 以片段化的 mRNA 为模板, 随机寡核苷酸为引物, 在 M-MuLV 逆转录酶体系中合成 cDNA 第 1 条链; 以 cDNA 第 1 条链为模板, 在 DNA 聚合酶 I、RNase H、dNTP 和缓冲液体系下合成 cDNA 第 2

链。用 1.8×Agencourt AMPure XP Beads 试剂盒 (Qiagen, Venlo, the Netherlands) 纯化双链 cDNA, 经末端修复和 3'非翻译区域(untranslated region, UTR)区加上 A 碱基后连接到 Illumina 测序接头上,使用 NEB # 7530 试剂盒 (New England Biolabs, CA, USA) 构建 cDNA 文库,文库构建 PCR 反应

体系为:纯化后双链 cDNA 模板 5 μL , 2×NEBNext High-Fidelity PCR Master Mix 25 μL , 通用 PCR 引物和 Index (X)引物各 1 μL , ddH₂O 补足至 50 μL ; 文库构建 PCR 反应程序为:98 °C 预变性 30 s;98 °C 变性 10 s, 65 °C 退火 75 s, 12 个循环。

表 1 基础饲料组成及营养水平(干物质基础)

Table 1 Compositions and nutrient levels of basal diets (DM basis)

| 项目 Items | 体质量阶段/kg Weight stage | | | |
|--|-----------------------|---------|---------|---------|
| | 150~250 | 250~350 | 350~450 | 450~550 |
| 蒸汽压片玉米/(g·kg ⁻¹) Steam-flaked corn | 180.0 | 218.4 | 400.0 | 610.0 |
| 蒸汽压片小麦/(g·kg ⁻¹) Steam-flaked wheat | 56.7 | 40.0 | 55.9 | 20.0 |
| 蒸汽压片大麦/(g·kg ⁻¹) Steam-flaked barley | 60.0 | 40.0 | 36.1 | 20.0 |
| 蒸汽压片黑豆/(g·kg ⁻¹) Steam-flaked black bean | 40.0 | 47.7 | 80.0 | 47.2 |
| 玉米胚芽粕/(g·kg ⁻¹) Corn germ meal | 60.0 | 48.2 | 20.0 | 60.0 |
| 玉米干酒糟及其可溶物/(g·kg ⁻¹) Corn DDGS | 130.6 | 160.0 | 31.9 | 30.0 |
| 碳酸氢钠/(g·kg ⁻¹) NaHCO ₃ | 5.0 | 6.7 | 7.4 | 9.4 |
| 食盐/(g·kg ⁻¹) NaCl | 5.0 | 6.7 | 7.4 | 9.4 |
| 石粉/(g·kg ⁻¹) Limestone | 5.0 | 9.0 | 6.3 | 5.0 |
| 预混料/(g·kg ⁻¹) Premix | 10.0 | 13.3 | 13.4 | 19.0 |
| 玉米黄贮/(g·kg ⁻¹) Corn stalk silage | 447.5 | 410.0 | 340.0 | 170.0 |
| 粗蛋白质/(g·kg ⁻¹) Crude protein | 140.0 | 126.0 | 113.0 | 100.9 |
| 维持净能/(KJ·kg ⁻¹) Net energy for maintenance | 4 530.0 | 6 290.0 | 7 750.0 | 8 580.0 |
| 增重净能/(KJ·kg ⁻¹) Net energy for gain | 3 700.0 | 4 700.0 | 5 130.0 | 5 880.0 |
| 钙/(g·kg ⁻¹) Calcium | 5.3 | 4.0 | 3.6 | 2.7 |
| 磷/(g·kg ⁻¹) Phosphorus | 3.0 | 3.9 | 3.1 | 2.9 |
| 中性洗涤纤维/(g·kg ⁻¹) Neutral detergent fiber | 312.1 | 293.1 | 226.3 | 158.7 |
| 酸性洗涤纤维/(g·kg ⁻¹) Acid detergent fiber | 179.6 | 160.2 | 129.0 | 78.7 |

注:每千克预混料含有:V_A 900 IU, V_D 150 IU, V_E 25 IU, Cu(硫酸铜)10 mg, Fe(硫酸亚铁)80 mg, Mn(硫酸锰)20 mg, Zn(硫酸锌)40 mg, I(乙二胺二氢碘化物)1.00 mg, Se(亚硒酸钠)0.60 mg。维持净能和增重净能为基于 NRC (2003) 的计算值,其余营养指标为实测值。

Note: Each kilogram of premix contains: V_A 900 IU, V_D 150 IU, V_E 25 IU, Cu(copper sulfate)10 mg, Fe(ferrous sulfate)80 mg, Mn(manganese sulfate)20 mg, Zn(zinc sulfate)40 mg, I(ethylenediamine dihydroiodide) 1.00 mg, and Se(sodium selenite)0.60 mg. Net energy for maintenance and net weight for gain were calculated values based on NRC (2003), while the other nutritional indicators were measured values.

使用 DNA 1000 assay Kit 试剂盒 (Agilent, Santa Clara, CA, USA) 检测 cDNA 文库样品片段的长度和浓度,质检合格后,在基迪奥生物技术有限公司(中国广州)的 Illumina HiSeq 2500 系统上进行 RNA-seq 测序。

1.3 新转录本的发掘与功能解析

利用 Fastp v0.18.0 软件对测序获得的原始读段(raw reads)进行过滤,去除含有 10% 以上未知核苷酸的读段、包含接头的读段、含 ploy-N 读段以及低质量读段,计算过滤后读段(clean reads)中质量值大于等于 20 的碱基占总碱基百分比(Q20)和质量值大于等于 30 的碱基占总碱基百分比(Q30)以及 GC 含量。应用 Hisat v2.2.4 软件将获得的过滤后读段与牛参考基因组(版本号:Ensembl release 96 Bos taurus ARS-UCD1.2)序列进行同源性对比,统计单一对比率、多重对比率、内含子对比率、外

显子对比率和基因间区对比率。利用 String Tie v1.3.1 软件寻找新转录本,将本次测序结果中发现的但未被牛参考基因组注释的转录本序列,过滤掉少于 50 个核苷酸残基的过短序列和只包含单个外显子的序列即得新转录本。将测序获得的转录本与参考转录本序列进行比较,延长已注释转录本的 3'和 5'URT,从而优化已注释转录本的信息,统计 3'和 5'UTR 延伸基因数、3'和 5'UTR 同时延伸基因数以及正链、负链延伸基因数。利用 String Tie v1.3.1 软件计算新转录本中每百万片段中来自某一基因每千碱基长度的数目(fragments per kilobase of transcript sequence per millions mapped reads,FPKM),将其作为表达量的单位对数据进行归一化处理,通过对显著性 P 值校正后得到错误发现率(false discovery rate, FDR),再计算两组样品间 FPKM 的比值作为差异倍数(fold change, FC),以

FDR<0.05 且 $|\log_2(\text{FC})| > 1$ 作为显著差异表达新转录本的阈值,筛选显著差异表达的新转录本。使用在线工具 DAVID v6.7 (<https://david.ncifcrf.gov/>)对筛选到的新转录本进行基因本体(gene ontology, GO)和京都基因和基因组百科全书(kyoto encyclopedia of genes and genomes, KEGG)富集分析。

1.4 肉牛新转录本的验证

为验证转录组分析数据的可靠性,以甘油醛-3-磷酸脱氢酶(glyceraldehyde-3-phosphate dehydrogenase, GAPDH)基因作为内参基因,选取 *MSTRG.6095*、*MSTRG.4813*、*MSTRG.1221* 等 9

个显著差异表达的新转录本(表 2)进行实时荧光定量 PCR (RT-qPCR)验证。根据 RNA-seq 得到的新转录本序列,利用 Primer 5 软件设计 RT-qPCR 特异性引物(表 2),交由陕西中科羽瞳生物科技有限公司合成。RT-qPCR 反应体系(20 μL)为:cDNA 模板 2 μL , 2 \times SYBR Green Pro Premix 10 μL , 10 $\mu\text{mol/L}$ 上、下游引物各 0.8 μL , ddH₂O 补足至 20 μL 。RT-qPCR 反应程序:95 $^{\circ}\text{C}$ 预变性 30 s; 95 $^{\circ}\text{C}$ 变性 5 s, 60 $^{\circ}\text{C}$ 退火 30 s, 40 个循环。反应在 Roche-Light Cycler 480 II 型荧光定量 PCR 仪上进行,采用 $2^{-\Delta\Delta\text{Ct}}$ 方法计算基因相对表达量^[20]。基因表达绘图时,将相对表达量换算成以 2 为底对数的差异倍数。

表 2 试验所用引物信息

Table 2 Information of primers used in the experiment

| 基因 Gene | 引物序列(5'→3')Primer sequence (5'→3') | 产物长度/bp Product size |
|--------------------|---|-------------------------|
| <i>MSTRG.6095</i> | F:GGTTCCAGGTGGTTTCAG;R:GTTGTCCGCCAGTTTGT | 169 |
| <i>MSTRG.4813</i> | F:GCTCCCTGAACCTTCCTCC;R:TCACCGCTCCGATTCCTAC | 122 |
| <i>MSTRG.1221</i> | F:CTGGACTGGAGGAAGAAGAG;R:ACGGGACAGGGTAGATGC | 136 |
| <i>MSTRG.13781</i> | F:AGTGCCCGTTACATTCA;R:GCTCATCAAGTTTGCCCTACA | 158 |
| <i>MSTRG.5937</i> | F:GGCGTCCGCTACAATACC;R:CTCCTGAGCCGACAGAAAG | 218 |
| <i>MSTRG.5068</i> | F:TCTCCCTGCATCTCCACC;R:CATCACATTTATTGCGTGTG | 173 |
| <i>MSTRG.2484</i> | F:CAGTATGGCTGCTGTGGAGA;R:TGTGGCCTACGTGGGTTT | 198 |
| <i>MSTRG.241</i> | F:TCTTGGAAGTTTCTGATTGC;R:GGTTTCTGGATAGGGCTGTTA | 130 |
| <i>MSTRG.10558</i> | F:GCAGTGGGTGAGCAAGAG;R:GGGATGGTTTCAGGACGA | 193 |
| <i>GAPDH</i> | F:AGCAATGCCTCTGTACCAC;R:AAGCAGGGATGATGATGTTCTGG | 139 |

2 结果与分析

2.1 肉牛 cDNA 文库信息及比对分析

本研究共构建了 9 个 cDNA 文库,每组 3 个,

Q20 均在 97% 以上, Q30 均在 94% 以上(表 3), 碱基 A、T、C 和 G 平均占比均约为 25%, AT 与 GC 碱基的含量不存在分离现象。

表 3 试验肉牛基因组 cDNA 文库信息

Table 3 cDNA library information of experimental beef cattle genome

| cDNA 文库 cDNA library | 原始读段/bp Raw reads | 过滤后读段/bp Clean reads | Q20/% | Q30/% | 碱基 G+C 含量/% G+C base percentage |
|-------------------------|----------------------|-------------------------|-------|-------|------------------------------------|
| J1 | 11 121 370 500 | 11 057 824 860 | 97.99 | 94.43 | 55.65 |
| J2 | 10 490 280 600 | 10 428 503 812 | 98.00 | 94.43 | 55.05 |
| J3 | 9 914 330 400 | 9 857 549 934 | 98.11 | 94.63 | 56.09 |
| P1 | 8 918 022 300 | 8 867 046 985 | 97.99 | 94.45 | 55.23 |
| P2 | 11 092 241 100 | 11 025 565 526 | 98.04 | 94.54 | 55.96 |
| P3 | 9 000 374 100 | 8 948 020 904 | 98.05 | 94.57 | 55.70 |
| S1 | 11 162 900 700 | 11 094 863 452 | 98.10 | 94.69 | 56.48 |
| S2 | 10 519 659 900 | 10 463 145 391 | 98.10 | 94.67 | 56.18 |
| S3 | 11 447 335 500 | 11 386 531 926 | 98.16 | 94.86 | 56.23 |

利用 Hisat2 软件将获得的过滤后读段与牛参考基因组(版本号:Ensembl release 96 Bos taurus ARS-UCD1.2)序列进行比对,结果(表 4)发现,单一比对率为 93.10%~94.05%,多重比对率为 2.99%~4.41%,内含子对比率为 4.55%~6.53%,外显子对比率为 83.28%~87.59%,基因间区对比率为 7.86%~10.20%。

2.2 肉牛已注释基因结构分析及优化

不同测序平台具有一定的局限性,获得的序列片段可能不够完整和精确,所以本研究利用转录组测序结果对已注释基因的结构进行优化,若已注释基因边界之外的区域存在与转录组测序结果连续匹配的读段,则该基因的 UTR 可向上下游延伸,从而优化基因边界^[21]。本研究对 3 705 个转录本进行了

结构优化,结果见表 5。

表 4 试验肉牛基因组 cDNA 文库过滤后读段与参考基因组序列的比对分析

Table 4 Comparison analysis of clean reads and reference genome of experimental beef cattle genome cDNA library

| cDNA 文库 cDNA library | 单一对比率/% Uniquely mapped | 多重对比率/% Multiply mapped | 内含子对比率/% Intron mapped | 外显子对比率/% Exon mapped | 基因间区对比率/% Intergenic region mapped |
|-------------------------|----------------------------|----------------------------|---------------------------|-------------------------|---------------------------------------|
| J1 | 93.57 | 3.43 | 5.55 | 85.96 | 8.49 |
| J2 | 94.05 | 2.99 | 5.50 | 85.18 | 9.32 |
| J3 | 93.95 | 3.21 | 5.18 | 86.23 | 8.59 |
| P1 | 93.57 | 3.10 | 6.53 | 83.28 | 10.20 |
| P2 | 93.73 | 3.44 | 5.13 | 86.43 | 8.44 |
| P3 | 94.05 | 3.22 | 5.21 | 86.00 | 8.79 |
| S1 | 93.51 | 3.67 | 5.62 | 85.46 | 8.92 |
| S2 | 93.18 | 4.08 | 5.09 | 85.95 | 8.96 |
| S3 | 93.10 | 4.41 | 4.55 | 87.59 | 7.86 |

表 5 基于转录组测序结果的肉牛群体已注释基因在 5' 和 3' UTR 延伸的统计分析

Table 5 Statistical analysis of annotated genes in 5' and 3' UTR extensions in beef cattle populations based on transcriptome sequencing results

| 染色体编号 Chromosome code | 已知转录本数 No of known transcripts | 5' UTR 延伸 基因数 No of genes extended at 5' UTR | 3' UTR 延伸 基因数 No of genes extended at 3' UTR | 5' 和 3' UTR 同时延伸 基因数 No of genes extended at the 5' and 3' UTR at the same time | 正链延伸基因数 No of positive strand extension genes | 负链延伸基因数 No of negative strand extension genes |
|-----------------------------|--------------------------------------|--|--|---|--|--|
| 1 | 153 | 37 | 70 | 46 | 71 | 82 |
| 2 | 164 | 46 | 70 | 48 | 83 | 81 |
| 3 | 173 | 47 | 79 | 47 | 88 | 85 |
| 4 | 113 | 36 | 48 | 29 | 65 | 48 |
| 5 | 183 | 58 | 74 | 51 | 90 | 93 |
| 6 | 103 | 27 | 50 | 26 | 52 | 51 |
| 7 | 189 | 57 | 76 | 56 | 93 | 96 |
| 8 | 96 | 20 | 53 | 23 | 51 | 45 |
| 9 | 91 | 21 | 37 | 33 | 51 | 40 |
| 10 | 151 | 40 | 64 | 47 | 78 | 73 |
| 11 | 169 | 37 | 76 | 56 | 82 | 87 |
| 12 | 73 | 26 | 26 | 21 | 33 | 40 |
| 13 | 147 | 45 | 67 | 35 | 78 | 69 |
| 14 | 87 | 27 | 37 | 23 | 48 | 39 |
| 15 | 121 | 39 | 53 | 29 | 64 | 57 |
| 16 | 126 | 28 | 49 | 49 | 69 | 57 |
| 17 | 125 | 39 | 46 | 40 | 62 | 63 |
| 18 | 201 | 79 | 59 | 63 | 112 | 89 |
| 19 | 228 | 66 | 101 | 61 | 117 | 111 |
| 20 | 50 | 13 | 22 | 15 | 21 | 29 |
| 21 | 100 | 31 | 46 | 23 | 43 | 57 |
| 22 | 126 | 29 | 63 | 34 | 67 | 59 |
| 23 | 113 | 33 | 47 | 33 | 53 | 60 |
| 24 | 64 | 15 | 27 | 22 | 26 | 38 |
| 25 | 155 | 52 | 77 | 26 | 85 | 70 |
| 26 | 75 | 15 | 36 | 24 | 36 | 39 |
| 27 | 49 | 10 | 27 | 12 | 26 | 23 |
| 28 | 63 | 19 | 25 | 19 | 35 | 28 |
| 29 | 93 | 47 | 31 | 15 | 47 | 46 |
| 30 | 121 | 41 | 48 | 32 | 59 | 62 |
| 其他 Other | 3 | 1 | 2 | 0 | 3 | 0 |
| 总计 Total | 3 705 | 1 081 | 1 586 | 1 038 | 1 888 | 1 817 |

表 5 显示,5' UTR 延伸的基因数有 1 081 个,3' UTR 延伸的基因数有 1 586 个,5' 和 3' UTR 同时延伸的基因数有 1 038 个,在正链上延伸的基因

数有 1 888 个,在负链上延伸的基因数有 1 817 个。在不同的染色体上随机各选取 1 个基因,对其优化前后的位置进行统计,结果(表 6)显示,有些已注释

基因在正链的 UTR 延伸,有些基因在负链的 UTR 延伸。

表 6 基于转录组测序结果的肉牛群体部分已注释基因的结构优化结果

Table 6 Structural optimization of partially annotated genes in beef cattle populations based on transcriptome sequencing results

| 染色体编号 Extended area | 基因编号 No of Chromosome | 延伸链 Gene ID | 延伸区域 Extended chain | 优化前位置/bp Location before optimization | 优化后位置/bp Location after optimization |
|------------------------|--------------------------|----------------|------------------------|--|---|
| 1 | ENSBTAT00000003985 | + | 3'UTR | 2 827 691~2 915 981 | 2 827 691~2 919 207 |
| 2 | ENSBTAT00000007163 | - | 5'UTR | 4 300 416~4 399 011 | 4 297 512~4 399 011 |
| 3 | ENSBTAT00000004378 | + | 3'UTR | 117 716 651~117 735 517 | 117 716 651~117 741 171 |
| 4 | ENSBTAT00000002003 | - | 5'UTR | 80 952 446~81 075 991 | 80 952 446~81 077 111 |
| 5 | ENSBTAT00000003810 | + | 3'UTR | 43 762 572~43 817 060 | 43 762 572~43 827 517 |
| 6 | ENSBTAT00000018687 | - | 3'UTR | 111 699 989~112 153 362 | 111 698 920~112 153 362 |
| 7 | ENSBTAT00000002376 | + | 5'UTR | 4 105 828~4 120 006 | 4 105 373~4 120 006 |
| 8 | ENSBTAT00000002523 | - | 3'UTR | 88 755 737~88 789 416 | 88 753 667~88 789 416 |
| 9 | ENSBTAT00000046332 | - | 5'UTR | 69 540 273~69 583 721 | 69 540 273~69 583 933 |
| 10 | ENSBTAT00000001668 | + | 3'UTR | 25 880 971~25 925 727 | 25 880 971~25 926 100 |
| 11 | ENSBTAT00000008106 | + | 5'UTR | 11 326 475~11 369 731 | 11 326 325~11 369 731 |
| 12 | ENSBTAT000000020126 | - | 3'UTR | 82 296 498~82 322 844 | 82 295 034~82 322 844 |
| 13 | ENSBTAT00000001268 | + | 5'UTR | 74 640 691~74 643 705 | 74 640 451~74 643 705 |
| 14 | ENSBTAT00000018083 | + | 3'UTR | 56 385 156~56 556 667 | 56 385 156~56 557 201 |
| 15 | ENSBTAT00000006638 | + | 5'UTR | 56 473 004~56 530 219 | 56 472 958~56 530 219 |
| 16 | ENSBTAT00000006773 | + | 3'UTR | 71 591 831~71 713 585 | 71 591 831~71 719 974 |
| 17 | ENSBTAT00000007876 | - | 5'UTR | 18 582 256~18 600 421 | 18 582 256~18 600 624 |
| 18 | ENSBTAT00000008257 | + | 3'UTR | 19 131 365~19 199 820 | 19 199 820~19 199 820 |
| 19 | ENSBTAT00000015224 | - | 5'UTR | 42 015 614~42 041 251 | 42 015 614~42 042 853 |
| 20 | ENSBTAT00000028278 | - | 3'UTR | 17 842 584~18 052 859 | 17 841 978~18 052 859 |
| 21 | ENSBTAT00000006650 | - | 5'UTR | 56 210 637~56 245 304 | 56 210 637~56 245 525 |
| 22 | ENSBTAT00000010230 | + | 3'UTR | 16 946 297~16 957 323 | 16 946 297~16 957 409 |
| 23 | ENSBTAT00000016029 | + | 5'UTR | 39 461 718~39 465 516 | 39 461 658~39 465 516 |
| 24 | ENSBTAT00000009214 | + | 3'UTR | 35 553 464~35 564 144 | 35 553 464~35 564 895 |
| 25 | ENSBTAT00000006410 | + | 5'UTR | 2 871 276~2 888 333 | 2 871 250~2 888 333 |
| 26 | ENSBTAT00000017090 | - | 3'UTR | 23 740 410~23 813 189 | 23 739 169~23 813 189 |
| 27 | ENSBTAT00000078362 | + | 3'UTR | 34 063 676~34 126 284 | 34 063 676~34 127 722 |
| 28 | ENSBTAT00000000186 | + | 5'UTR | 28 287 394~28 297 656 | 28 286 557~28 297 656 |
| 29 | ENSBTAT00000076965 | + | 5'UTR | 43 758 770~43 774 081 | 43 758 480~43 774 081 |
| 30 | ENSBTAT00000012726 | + | 3'UTR | 33 366 140~33 425 170 | 33 366 140~33 426 878 |

注: +, 正链; -, 负链。

Note: +, Positive strand; -, Negative strand.

2.3 肉牛新转录本的发掘与功能注释

本研究共获得 442 个新的转录本,其长度为 215~21 315 bp,平均长度为 1 930 bp,不同长度转录本的分布见图 1。

利用 DAVID 软件对筛选到的新转录本进行 GO 富集分析,442 个新转录本注释到生物学过程、细胞组分和分子功能。在生物学过程中,细胞过程、单组织过程、代谢过程、生物调节等的基因较多;在分子功能中,绑定富集的基因最多,其次是酶催化活性和分子换能器活性等;在细胞组分中,细胞、细胞部分、细胞器、生物膜(膜和膜部分)等富集的基因较多(图 2)。KEGG 富集分析发现,442 个新转录本富集到 148 条通路,其中味觉转导、癌症通路以及河马(Hippo)信号通路等途径富集的基因较多(图 3)。

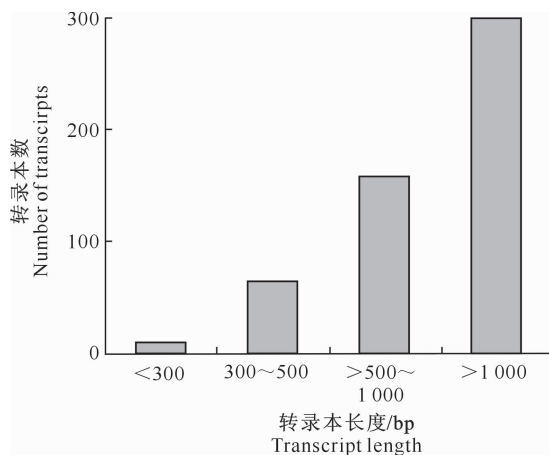
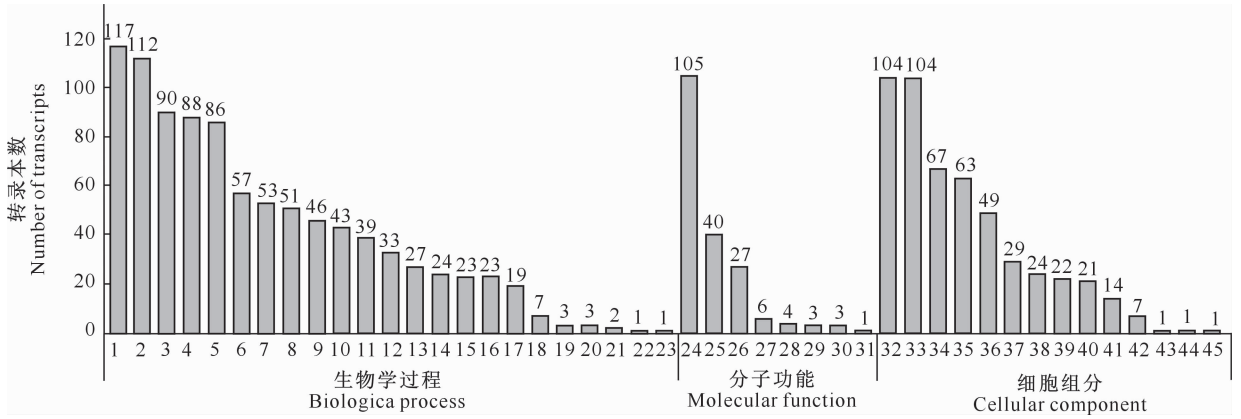


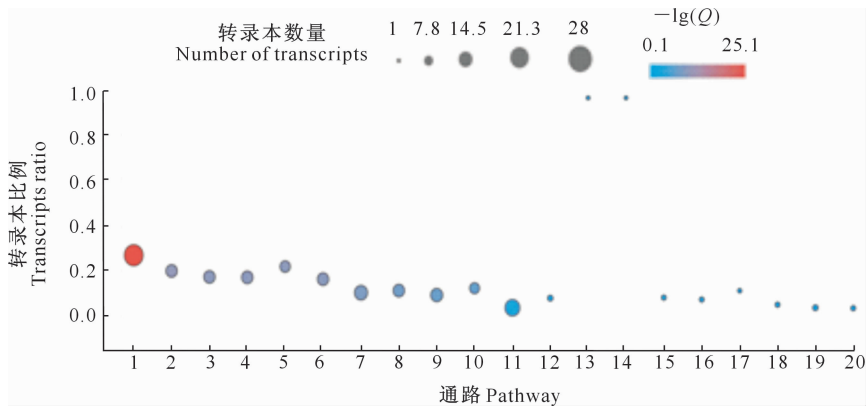
图 1 3 个肉牛群体新转录本的长度分布
Fig. 1 Distribution of new transcript lengths among 3 beef cattle populations



1. 细胞过程; 2. 单组织过程; 3. 代谢过程; 4. 生物调节; 5. 生物过程的调节; 6. 刺激反应; 7. 多细胞生物过程; 8. 位置; 9. 信号; 10. 发育过程; 11. 生物过程的负调节; 12. 细胞成分组织或生物生成; 13. 生物附着; 14. 免疫系统过程; 15. 生物过程的正调节; 16. 运动; 17. 行为; 18. 多生物过程; 19. 繁殖; 20. 生殖过程; 21. 细胞死亡; 22. 生长; 23. 节奏过程; 24. 绑定; 25. 酶催化活性; 26. 分子换能器活性; 27. 运输活动; 28. 核酸结合转录因子活性; 29. 信号传感器活动; 30. 结构分子活性; 31. 转录因子活性, 蛋白质结合; 32. 细胞; 33. 细胞部分; 34. 细胞器; 35. 膜; 36. 膜部分; 37. 细胞器部分; 38. 细胞外区域; 39. 大分子复合物; 40. 细胞外区域部分; 41. 细胞连接; 42. 膜封闭腔; 43. 胞外基质; 44. 病毒部分; 45. 病毒
1. Cellular process; 2. Single-organism process; 3. Metabolic process; 4. Biological regulation; 5. Regulation of biological process; 6. Response to stimulus; 7. Multicellular organismal process; 8. Localization; 9. Signaling; 10. Developmental process; 11. Negative regulation of biological process; 12. Cellular component organization or biogenesis; 13. Biological adhesion; 14. Immune system process; 15. Positive regulation of biological process; 16. Locomotion; 17. Behavior; 18. Multi-organism process; 19. Reproduction; 20. Reproductive process; 21. Cell killing; 22. Growth; 23. Rhythmic process; 24. Binding; 25. Enzyme catalytic activity; 26. Molecular transducer activity; 27. Transporter activity; 28. Nucleic acid binding transcription factor activity; 29. Signal transducer activity; 30. Structural molecule activity; 31. Transcription factor activity, protein binding; 32. Cell; 33. Cell part; 34. Organelle; 35. Membrane; 36. Membrane part; 37. Organelle part; 38. Extracellular region; 39. Macromolecular complex; 40. Extracellular region part; 41. Cell junction; 42. Membrane-enclosed lumen; 43. Extracellular matrix; 44. Virion part; 45. Virion

图 2 3 个肉牛群体新转录本的 GO 注释

Fig. 2 GO annotation of new transcripts among 3 beef cattle populations



- Q. 经过校正后的 P 值; 1. 味觉传导; 2. 子宫内膜癌; 3. 致心律失常右室心肌病; 4. 黏着连接; 5. 核苷酸切除修复; 6. 上皮细胞受细菌侵袭; 7. 胃癌; 8. 白细胞跨内皮迁移; 9. 河马(Hippo)信号通路; 10. 过氧化物酶; 11. 癌症通路; 12. ABC 转运蛋白; 13. MAPK 信号通路-fly; 14. 光传导; 15. 抗叶酸; 16. 叶酸生物合成; 17. 初级胆汁酸生物合成; 18. 卵巢类固醇生成; 19. 糖尿病并发症中的 AGE-RAGE 信号通路; 20. 花生四烯酸代谢

- Q. Corrected P-value; 1. Taste transduction; 2. Endometrial cancer; 3. Arrhythmogenic right ventricular cardiomyopathy (ARVC); 4. Adherens junction; 5. Nucleotide excision repair; 6. Bacterial invasion of epithelial cells; 7. Gastric cancer; 8. Leukocyte transendothelial migration; 9. Hippo signaling pathway; 10. Peroxisome; 11. Pathways in cancer; 12. ABC transporters; 13. MAPK signaling pathway-fly; 14. Phototransduction-fly; 15. Antifolate resistance; 16. Folate biosynthesis; 17. Primary bile acid biosynthesis; 18. Ovarian Steroidogenesis; 19. AGE-RAGE signaling pathway in diabetic complications; 20. Arachidonic acid metabolism

图 3 3 个肉牛群体新转录本的 KEGG 分析

Fig. 3 KEGG analysis of new transcripts among 3 beef cattle populations

2.4 肉牛新转录本差异表达分析

达分析结果(表 7~9)发现,共有 92 个在不同群体

在不同肉牛群体的 442 个新转录本中,差异表 肉牛中显著差异表达($P < 0.05$)。

表 7 杂交和牛和西门塔尔牛组间显著差异表达的新转录本

Table 7 New transcripts with significantly differential expression between crossbred Wagyu and Simmental cattle groups

| 转录本名称 Transcript ID | lb(FC) | FDR | 转录本名称 Transcript ID | lb(FC) | FDR | 转录本名称 Transcript ID | lb(FC) | FDR |
|------------------------|--------|--------|------------------------|--------|--------|------------------------|--------|--------|
| MSTRG. 1221 | 1.05 | <0.001 | MSTRG. 10117 | 3.56 | <0.001 | MSTRG. 242 | -1.41 | 0.003 |
| MSTRG. 6095 | 1.71 | <0.001 | MSTRG. 574 | 3.41 | 0.049 | MSTRG. 3862 | -3.32 | <0.001 |
| MSTRG. 4813 | 3.52 | <0.001 | MSTRG. 9413 | 2.73 | <0.001 | MSTRG. 10683 | -2.00 | <0.001 |
| MSTRG. 215 | 3.32 | 0.020 | MSTRG. 2110 | 4.34 | 0.043 | MSTRG. 440 | -1.78 | <0.001 |
| MSTRG. 10230 | 2.34 | <0.001 | MSTRG. 7837 | 3.39 | 0.016 | MSTRG. 13399 | -1.84 | 0.017 |
| MSTRG. 3577 | 1.25 | 0.027 | MSTRG. 2377 | 3.33 | 0.017 | MSTRG. 10690 | -3.01 | <0.001 |
| MSTRG. 9856 | 1.95 | 0.028 | MSTRG. 5068 | -2.23 | <0.001 | MSTRG. 10225 | -1.56 | 0.022 |
| MSTRG. 14513 | 10.11 | <0.001 | MSTRG. 9976 | -1.10 | <0.001 | MSTRG. 6493 | -1.50 | 0.002 |
| MSTRG. 3086 | 5.21 | 0.037 | MSTRG. 13812 | -2.16 | <0.001 | MSTRG. 533 | -2.31 | 0.004 |
| MSTRG. 14512 | 9.96 | <0.001 | MSTRG. 6263 | -1.02 | <0.001 | MSTRG. 14169 | -2.18 | 0.031 |
| MSTRG. 4549 | 1.47 | 0.043 | MSTRG. 9073 | -1.11 | 0.007 | MSTRG. 6849 | -3.10 | <0.001 |
| MSTRG. 10607 | 1.53 | 0.016 | MSTRG. 10558 | -1.53 | <0.001 | MSTRG. 11123 | -3.28 | 0.012 |
| MSTRG. 14719 | 1.16 | 0.017 | MSTRG. 1182 | -1.58 | <0.001 | MSTRG. 11843 | -3.02 | 0.024 |
| MSTRG. 7766 | 1.11 | 0.005 | MSTRG. 5937 | -2.93 | <0.001 | MSTRG. 1401 | -3.22 | 0.011 |
| MSTRG. 8941 | 1.38 | 0.003 | MSTRG. 13781 | -1.94 | <0.001 | MSTRG. 3522 | -4.11 | 0.018 |
| MSTRG. 12300 | 8.08 | 0.007 | MSTRG. 12167 | -1.04 | 0.002 | MSTRG. 922 | -3.95 | 0.029 |
| MSTRG. 383 | 2.36 | 0.008 | MSTRG. 1686 | -1.50 | 0.005 | MSTRG. 11523 | -10.05 | <0.001 |

注:FC为杂交和牛和西门塔尔牛新转录本表达量的比值;正值表示杂交和牛较西门塔尔牛显著上调,负值表示杂交和牛较西门塔尔牛显著下调。

Note:FC is the ratio of expression of new transcripts between crossbred Wagyu and Simmental cattle;Positive values indicate significant up-regulation of crossbred Wagyu compared to Simmental cattle,and negative values indicate significant down-regulation.

表 8 平凉红牛和西门塔尔牛组间显著差异表达的新转录本

Table 8 New transcripts with significantly differential expression between Pingliang red cattle and Simmental cattle groups

| 转录本名称 Transcript ID | lb(FC) | FDR | 转录本名称 Transcript ID | lb(FC) | FDR | 转录本名称 Transcript ID | lb(FC) | FDR |
|------------------------|--------|--------|------------------------|--------|--------|------------------------|--------|--------|
| MSTRG. 375 | 1.06 | 0.006 | MSTRG. 4813 | 2.32 | <0.001 | MSTRG. 10690 | -2.52 | <0.001 |
| MSTRG. 1221 | 1.11 | 0.001 | MSTRG. 215 | 4.12 | <0.001 | MSTRG. 5937 | -2.09 | <0.001 |
| MSTRG. 13383 | 1.13 | 0.012 | MSTRG. 3086 | 4.81 | 0.018 | MSTRG. 1182 | -1.67 | <0.001 |
| MSTRG. 10230 | 1.55 | <0.001 | MSTRG. 12300 | 5.70 | 0.011 | MSTRG. 6849 | -1.66 | <0.001 |
| MSTRG. 10607 | 1.68 | 0.007 | MSTRG. 11620 | 5.78 | 0.015 | MSTRG. 10683 | -1.49 | 0.005 |
| MSTRG. 11512 | 1.82 | 0.019 | MSTRG. 11523 | -5.72 | <0.001 | MSTRG. 6007 | -1.37 | 0.023 |
| MSTRG. 9098 | 1.90 | 0.003 | MSTRG. 922 | -4.78 | 0.028 | MSTRG. 5068 | -1.34 | <0.001 |
| MSTRG. 6095 | 1.94 | <0.001 | MSTRG. 14169 | -3.03 | <0.001 | MSTRG. 10558 | -1.27 | <0.001 |
| MSTRG. 3926 | 1.98 | <0.001 | MSTRG. 533 | -2.88 | <0.001 | MSTRG. 13781 | -1.19 | <0.001 |
| MSTRG. 383 | 2.05 | 0.022 | MSTRG. 3862 | -2.71 | <0.001 | MSTRG. 13812 | -1.17 | <0.001 |

注:FC为平凉红牛和西门塔尔牛新转录本表达量的比值。正值表示平凉红牛较西门塔尔牛显著上调,负值表示平凉红牛较西门塔尔牛显著下调。

Note:FC is the ratio of expression of new transcripts between Pingliang red cattle and Simmental cattle. Positive values indicate significant up-regulation of Pingliang red cattle compared to Simmental cattle,and negative values indicate significant down-regulation.

表 9 杂交和牛和平凉红牛组间显著差异表达的新转录本

Table 9 New transcripts with significantly differential expression between crossbred Wagyu and Pingliang red cattle groups

| 转录本名称 Transcript ID | lb(FC) | FDR | 转录本名称 Transcript ID | lb(FC) | FDR | 转录本名称 Transcript ID | lb(FC) | FDR |
|------------------------|--------|--------|------------------------|--------|--------|------------------------|--------|-------|
| MSTRG. 8014 | 9.75 | <0.001 | MSTRG. 8448 | -2.19 | 0.001 | MSTRG. 3519 | -3.01 | 0.015 |
| MSTRG. 4813 | 1.20 | <0.001 | MSTRG. 3521 | -2.88 | <0.001 | MSTRG. 3522 | -4.44 | 0.039 |
| MSTRG. 14512 | 9.96 | 0.005 | MSTRG. 2484 | -1.29 | <0.001 | MSTRG. 241 | -1.05 | 0.049 |
| MSTRG. 14223 | -6.90 | <0.001 | MSTRG. 2137 | -2.73 | <0.001 | | | |

注:FC为杂交和牛和平凉红牛新转录本表达量的比值;正值表示杂交和牛较平凉红牛显著上调,负值表示杂交和牛较平凉红牛显著下调。

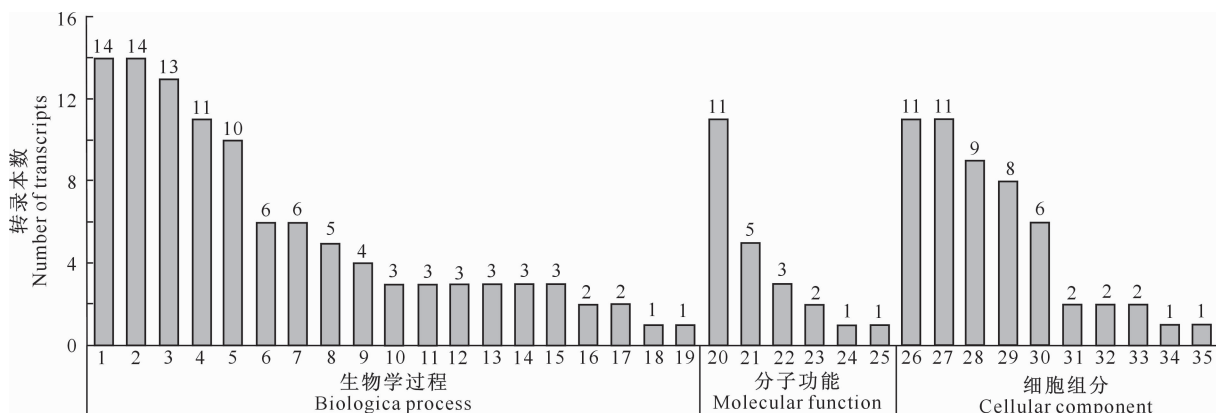
Note:FC is the ratio of expression of new transcripts between crossbred Wagyu and Pingliang red cattle. Positive values indicate significant up-regulation of crossbred Wagyu compared to Pingliang red cattle and negative values indicate significant down-regulation.

由表 7~9 可知,在 92 个显著差异表达的转录本中,杂交和牛比西门塔尔牛显著上调 23 个,显著下调 28 个;平凉红牛比西门塔尔牛显著上调和显著下调各 15 个;杂交和牛比平凉红牛显著上调 3 个,显著下调 8 个。

GO 富集分析表明,在生物学过程中,显著差异

表达的新转录本在单组织过程、细胞过程和代谢过程的富集程度最高;在分子功能中,绑定、分子换能器活性和酶催化活性的富集程度最高;在细胞组分中,细胞和细胞部分富集程度最高(图 4~6)。

3 个肉牛群体显著差异表达新转录本的 KEGG 注释如图 7 所示。



1. 单组织过程; 2. 代谢过程; 3. 细胞过程; 4. 生物调节; 5. 生物过程的调节; 6. 刺激反应; 7. 多细胞生物过程;
8. 信号; 9. 发育过程; 10. 行为; 11. 位置; 12. 生物过程的负调节; 13. 生物附着; 14. 运动; 15. 免疫系统过程;
16. 细胞成分组织或生物生成; 17. 生物过程的正调节; 18. 繁殖; 19. 生殖过程; 20. 绑定; 21. 分子换能器活性; 22. 酶催化活性;
23. 信号传感器活动; 24. 核酸结合转录因子活性; 25. 转录因子活性, 蛋白质结合; 26. 细胞; 27. 细胞部分; 28. 膜; 29. 膜部分;
30. 细胞器; 31. 大分子复合物; 32. 细胞外区域部分; 33. 细胞连接; 34. 细胞器部分; 35. 细胞外区域部分; 36. 胞外基质。图 5 和图 6 同
1. Single-organism process; 2. Metabolic process; 3. Cellular process; 4. Biological regulation; 5. Regulation of biological process;
6. Response to stimulus; 7. Multicellular organismal process; 8. signaling; 9. Developmental process; 10. Behavior; 11. Localization;
12. Negative regulation of biological process; 13. Biological adhesion; 14. Locomotion; 15. Immune system process;
16. Cellular component organization or biogenesis; 17. Positive regulation of biological process; 18. Reproduction; 19. Reproductive process;
20. Binding; 21. Molecular transducer activity; 22. Enzyme catalytic activity; 23. Signal transducer activity; 24. Nucleic acid binding transcription factor activity; 25. Transcription factor activity, protein binding; 26. Cell; 27. Cell part; 28. Membrane; 29. Membrane part;
30. Organelle; 31. Macromolecular complex; 32. Extracellular region; 33. Cell junction; 34. Organelle part; 35. Extracellular region part;
36. Extracellular matrix; Fig. 5 and Fig. 6 are same

图 4 杂交和牛和西门塔尔牛差异新转录本 GO 注释

Fig. 4 GO annotation of differential new transcripts between crossbred Wagyu and Simmental cattle

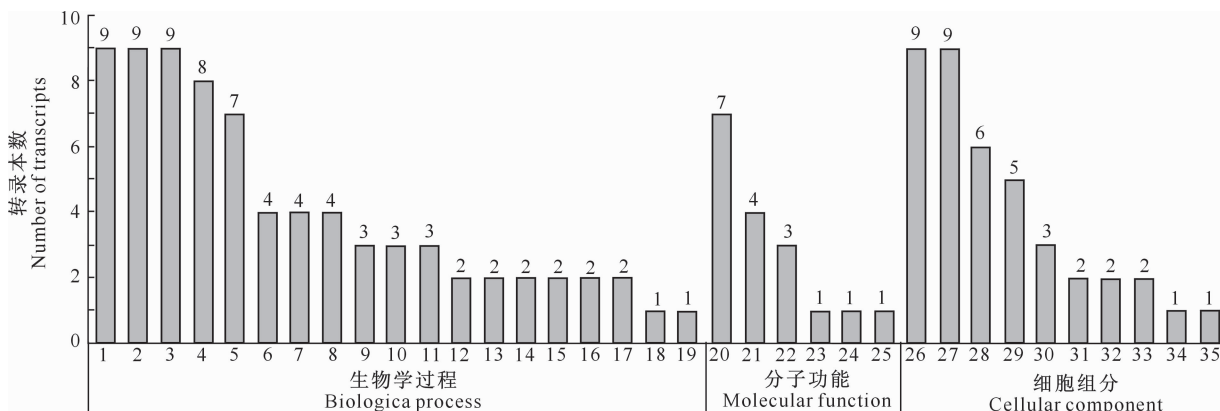


图 5 平凉红牛和西门塔尔牛差异新转录本 GO 注释

Fig. 5 GO annotation of differential new transcripts between Pingliang red cattle and Simmental cattle

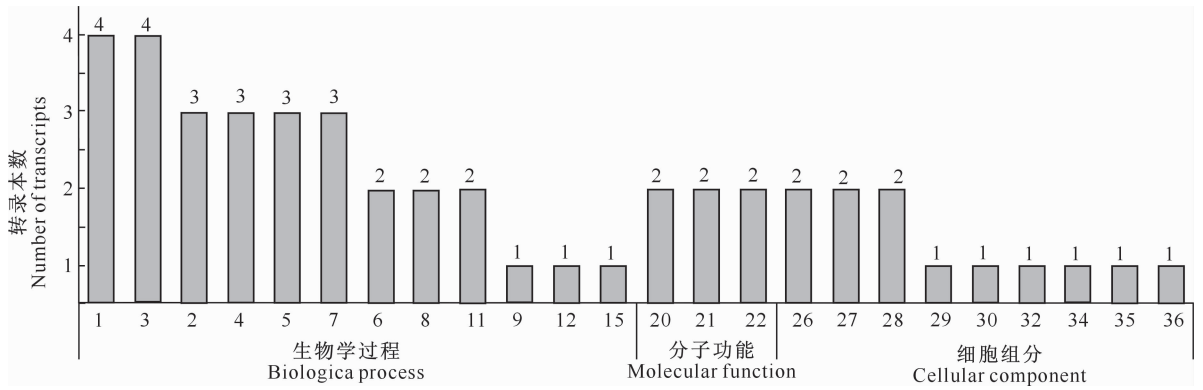
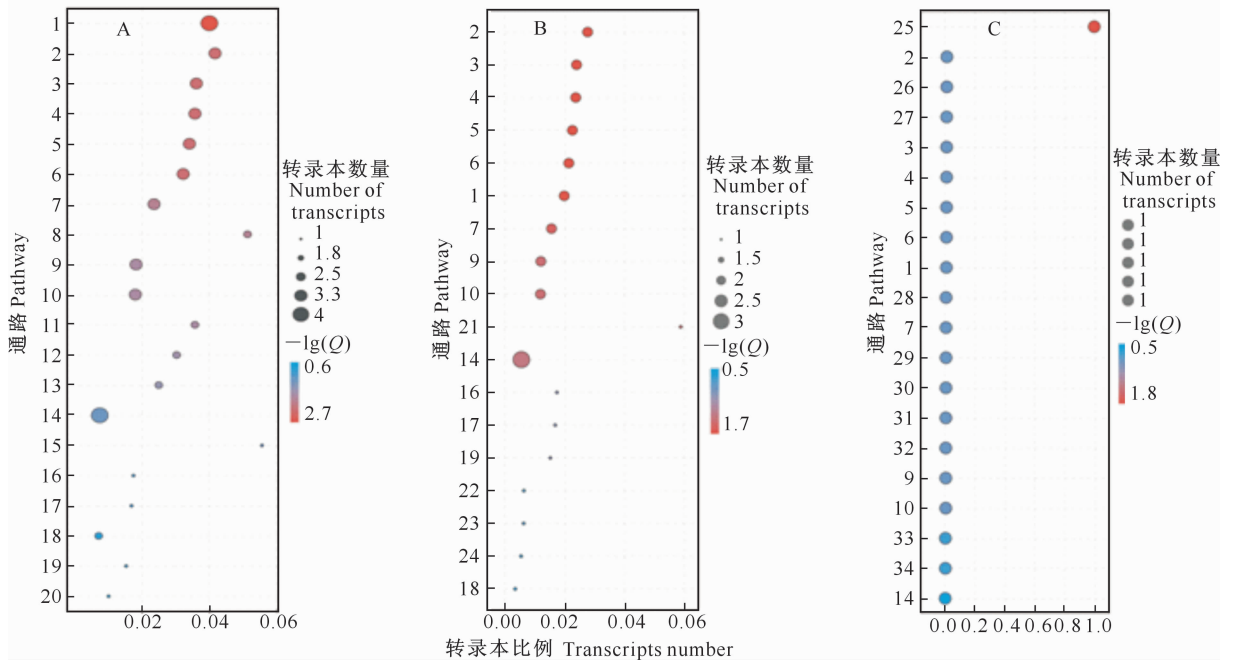


图 6 平凉红牛和杂交和牛差异新转录本 GO 注释

Fig. 6 GO annotation of differential new transcripts between Pingliang red cattle and crossbred Wagyu



- A. 杂交和牛和西门塔尔牛; B. 平凉红牛和西门塔尔牛; C. 平凉红牛和杂交和牛。Q. 经过校正后的 P 值。1. 味觉转导; 2. 子宫内膜癌; 3. 致心律失常右心室心肌病; 4. 黏着连接; 5. 上皮细胞受细菌侵袭; 6. 过氧化物酶; 7. 白细胞跨内皮迁移; 8. 叶酸生物合成; 9. 河马 (Hippo) 信号通路; 10. 胃癌; 11. 卵巢类固醇生成; 12. 类固醇激素生物合成; 13. 花生四烯酸代谢; 14. 癌症通路; 15. 氮代谢; 16. 核苷酸切除修复; 17. 长期抑郁; 18. 神经活性配体-受体相互作用; 19. 急性髓系白血病; 20. 炎症介质对色氨酸通道的调节; 21. 初级胆汁酸生物合成; 22. 库欣综合征; 23. 非酒精性脂肪肝; 24. 癌症中的转录失调; 25. MAPK 信号通路-fly; 26. p53 信号通路; 27. 唾液分泌; 28. 自然杀伤细胞介导的细胞毒性; 29. 麻疹; 30. 细胞凋亡; 31. 坏死性凋亡; 32. 甲型流感; 33. 细胞因子-细胞因子受体相互作用; 34. MAPK 信号通路

A. Crossbred Wagyu and Simmental cattle; B. Pingliang red cattle and Simmental cattle; C. Pingliang red cattle and crossbred Wagyu.

- Q. Corrected P-value. 1. Taste transduction; 2. Endometrial cancer; 3. Arrhythmogenic right ventricular cardiomyopathy (ARVC); 4. Adherens junction; 5. Bacterial invasion of epithelial cells; 6. Peroxisome; 7. Leukocyte transendothelial migration; 8. Folate biosynthesis; 9. Hippo signaling pathway; 10. Gastric cancer; 11. Ovarian steroidogenesis; 12. Steroid hormone biosynthesis; 13. Arachidonic acid metabolism; 14. Pathways in cancer; 15. Nitrogen metabolism; 16. Nucleotide excision repair; 17. Long-term depression; 18. Neuroactive ligand-receptor interaction; 19. Acute myeloid leukemia; 20. Inflammatory mediator regulation of TRP channels; 21. Primary bile acid biosynthesis; 22. Cushing syndrome; 23. Non-alcoholic fatty liver disease (NAFLD); 24. Transcriptional misregulation in cancers; 25. MAPK signaling pathway-fly; 26. p53 signaling pathway; 27. Salivary secretion; 28. Natural killer cell mediated cytotoxicity; 29. Measles; 30. Apoptosis; 31. Necroptosis; 32. Influenza A; 33. Cytokine-cytokine receptor interaction; 34. MAPK signaling pathway

图 7 3 个肉牛群体显著差异表达新转录本的 KEGG 注释

Fig. 7 KEGG annotation of significantly differentially expressing new transcripts among 3 beef cattle

KEGG 功能富集结果(图 7)显示,杂交和牛和西门塔尔牛显著差异表达的新转录本主要富集于味觉转导、癌症通路、过氧化物酶和河马(Hippo)信号通路等通路,平凉红牛和西门塔尔牛显著差异表达的新转录本主要富集于癌症通路、河马(Hippo)信号通路和味觉转导通路,而杂交和牛和平凉红牛显

著差异表达的新转录本主要富集于 MAPK 和 p53 (p53 signaling pathway)等信号通路。随机挑选 9 个新转录本进行 RT-qPCR 验证,结果(图 8)显示,RT-qPCR 结果与 RNA-seq 结果在定量方面存在差异,但变化趋势一致,说明转录组测序结果可靠。

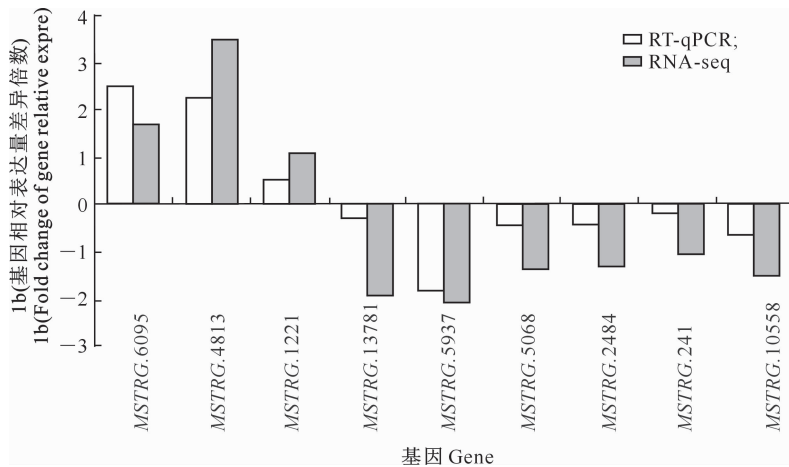


图 8 3 个肉牛群体部分显著差异表达基因的 RT-qPCR 验证

Fig. 8 RT-qPCR validation of partial genes with significantly differential expression among 3 beef cattle populations

3 讨论

3.1 牛基因组结构及新转录本

RNA-seq 技术在完善基因组结构信息和发掘新转录本方面有重要作用,已有研究人员利用该技术在牦牛^[22]和绵羊^[23]上分别挖掘出 11 362 和 9 001 个新转录本。牛全基因组测序、组装和注释已完成^[24],但基因序列和功能注释并不完善^[25],因此本研究利用测序获得的 3 个肉牛群体的背最长肌转录组数据,对牛基因组的基因序列信息进行了结构优化,共优化了 3 705 个已注释转录本,其中 5'UTR 延伸的转录本有 1 081 个,3'UTR 延伸的转录本有 1 586 个,5'和 3'UTR 同时延伸的转录本有 1 038 个;在正链上延伸的转录本有 1 888 个,在负链上延伸的转录本有 1 817 个。韩信嵘^[6]在西门塔尔牛和安格斯牛背最长肌中发现了 993 个新转录本,本研究在平凉红牛、杂交和牛和西门塔尔牛中挖掘出 442 个新转录本,这些新转录本富集到 148 条通路上。本研究结果丰富和完善了牛的转录本数据库,为基因功能注释提供了参考依据。

3.2 GO 和 KEGG 富集分析

肌内脂肪是影响牛肉品质的重要因素,本课题组前期的研究发现,平凉红牛肌内脂肪和其他肉品质指标基本达到杂交和牛的水平,明显优于西门塔

尔牛^[26]。本研究通过对新转录本的 GO 富集分析发现,大多新转录本富集在细胞过程、单组织过程、代谢过程、生物调节等生物学过程。一般来说,代谢通路上的关键调节酶活性能够反映代谢过程的强度和活性,如无氧代谢通路中的丙酮酸激酶为细胞提供乙酰辅酶 A 作为脂质生物合成(脂肪酸和胆固醇)的前体^[27],无氧代谢增强时,肌肉中乳酸含量增多进而影响肉品质^[28]。本研究在牛背最长肌中发现,具有酶催化活性功能的基因显著富集,可能这些基因通过无氧代谢酶调控机体的乳酸含量和脂质生物合成途径,进而改善肉的品质。

有研究发现,味觉转导通路与胃肠道功能、食物摄入和葡萄糖代谢有紧密联系^[29],河马(Hippo)信号通路在器官发育、组织稳态和细胞再生方面发挥着关键作用^[30],过氧化物酶参与多种代谢过程,包括哺乳动物体内长链脂肪酸的 β 氧化,胆汁酸和二十二碳六烯酸(DHA)的合成^[31]。本试验对差异新转录本的 KEGG 富集结果发现,杂交和牛和西门塔尔牛差异新转录本,显著富集于味觉转导、Hippo 信号通路以及过氧化物酶等途径,平凉红牛和西门塔尔牛差异新转录本显著富集于癌症通路、Hippo 信号通路和味觉转导等。由此可见,平凉红牛与杂交和牛之间的差异新转录本可能在牛发育过程中具有重要调节作用,同时广泛参与调控牛机体营养摄取

和脂质代谢。本研究中,平凉红牛和杂交和牛差异新转录本显著富集于 MAPK 和 p53 等信号通路。MAPK 信号通路参与牛肌内脂肪沉积的调控^[32],对于猪肌内脂肪细胞而言,CTRP6 基因通过 MAPK 信号通路促进其分化^[33];p53 信号通路除了具有众所周知的抑癌功能外,还在脂肪组织的代谢中发挥重要作用^[34];CTRP6 基因被敲除后,MAPK、p53 和脂肪因子信号通路在猪皮下脂肪细胞中特异性富集^[35]。因此推测本研究中富集在 MAPK 和 p53 等信号通路上的差异新转录本,有可能参与牛肌内脂肪沉积的调控。另外,本研究还发现了大量富集于癌症通路的新转录本。这些新转录本的发现既为揭示牛基因组信息及其功能提供了参考,也为进一步了解脂肪沉积与肉质差异的遗传机制提供了基础资料。

4 结 论

共挖掘出肉牛新转录本 442 个,其中显著差异表达的新转录本有 92 个。新转录本富集较多的条目均与细胞功能和代谢相关;KEGG 富集的显著差异表达的新转录本主要在味觉转导、过氧化物酶、河马(Hippo)信号通路和 MAPK 等信号通路中。这些新转录本通过各种代谢途径和通路,广泛参与调控牛肌肉生长、脂肪沉积、疾病和免疫等各项生命活动。

[参考文献]

- [1] Liu T, Wu J P, Lei Z M, et al. Fatty acid profile of muscles from crossbred angus-simmental, wagyu-simmental, and chinese simmental cattles [J]. *Food Sci Anim Resour*, 2020, 40(4): 563-577.
- [2] Gotoh T, Nishimura T, Kuchida K, et al. The japanese wagyu beef industry: current situation and future prospects: a review [J]. *Asian Austral J Anim*, 2018, 31(7): 933-950.
- [3] 李荣生. “平凉红牛”产业发展问题研究 [J]. *中国牛业科学*, 2018, 44(1): 64-69.
Li R S. Research on “pingliang red cattle” industrial developrnt [J]. *China Cattle Science*, 2018, 44(1): 64-69.
- [4] 马 钧, 张路培, 陈 燕, 等. 平凉红牛群体特性研究进展及未来育种研究方向 [J]. *中国畜牧杂志*, 2021, 57(11): 1-8.
Ma J, Zhang L P, Chen Y, et al. Research progress on population characteristics and future breeding research directions of pingliang red cattle [J]. *Chinese Journal of Animal Science*, 2021, 57(11): 1-8.
- [5] 马志锋, 曹效锋, 李新科, 等. 日本和牛冻精杂交本地“平凉红牛”效果 [J]. *畜牧兽医杂志*, 2017, 36(4): 12-14.
Ma Z F, Cao X F, Li X K, et al. Hybridization effect of Pingliang red bull and Japanese Wagyu (frozen semen) [J]. *Journal of Animal Science and Veterinary Medicine*, 2017, 36(4): 12-14.
- [6] 韩信嵘. 基于转录组测序分析中国西门塔尔牛和安格斯牛脂肪沉积相关基因的差异表达 [D]. 内蒙古通辽: 内蒙古民族大学, 2020.
Han X R. Analysis of differentially expressed genes related to fat deposition in Simmental and Angus based on transcriptome sequencing [D]. Tongliao, Inner Mongolia: Inner mongolia University for Nationalities, 2020.
- [7] Mirza B, Wang W, Wang J, et al. Machine learning and integrative analysis of biomedical big data genes [J]. *Basel*, 2019, 10(2): 87.
- [8] 裴晓婷, 焦新伟, 平智广. 转录组数据的图形呈现 [J]. *中华预防医学杂志*, 2020, 54(5): 586-592.
Pei X T, Jiao X W, Ping Z G. Visualization of transcriptome data [J]. *Chinese Journal of Preventive Medicine*, 2020, 54(5): 586-592.
- [9] Zhao Y, Wang K, Wang W L, et al. A high-throughput SNP discovery strategy for RNA-seq data [J]. *BMC Genomics*, 2019, 20(1): 160.
- [10] Owens N D L, De Domenico E, Gilchrist M J. An RNA-Seq protocol for differential expression analysis [J]. *Cold Spring Harb Protoc*, 2019(6).
- [11] Wang B, Wang P, Parobchak N, et al. Integrated RNA-seq and ChIP-seq analysis reveals a feed-forward loop regulating H3K9ac and key labor drivers in human placenta [J]. *Placenta*, 2019, 76: 40-50.
- [12] Holland C H, Tanevski J, Perales-Patón J, et al. Robustness and applicability of transcription factor and pathway analysis tools on single-cell RNA-seq data [J]. *Genome Biol*, 2020, 21(1): 36.
- [13] Xing K, Liu H, Zhang F, et al. Identification of key genes affecting porcine fat deposition based on co-expression network analysis of weighted genes [J]. *J Anim Sci Biotechnol*, 2021, 12(1): 100.
- [14] Zhang J S, Xu H Y, Fang J C, et al. Integrated microRNA-mRNA analysis reveals the roles of microRNAs in the muscle fat metabolism of Yanbian cattle [J]. *Anim Genet*, 2021, 52(5): 598-607.
- [15] Chen X, Luo Y, Wang R, et al. Effects of fatty acid transport protein 1 on proliferation and differentiation of porcine intramuscular preadipocytes [J]. *Anim Sci J*, 2017, 88(5): 731-738.
- [16] Rizzo M R, Fasano R, Paolisso G. Adiponectin and cognitive decline [J]. *Int J Mol Sci*, 2020, 21(6).
- [17] 郭 睿, 张 璐, 熊翠玲, 等. 西方蜜蜂的基因注释信息优化及新基因鉴定 [J]. *上海交通大学学报(农业科学版)*, 2018, 36(3): 39-44.
Guo R, Zhang L, Xiong C L, et al. Optimization of annotated genes' information and identification of novel genes in *Apis mellifera* [J]. *Journal of Shanghai Jiaotong University(Agri-*

- cultural Science), 2018, 36(3): 39-44.
- [18] 郭英飞, 王玉飞, 龚春丽, 等. 基于 RNA-Seq 的羊种布鲁氏菌新转录本与非编码 RNA 鉴定 [J]. 中国人兽共患病学报, 2015, 31(3): 216-221.
Guo Y F, Wang Y F, Gong C L, et al. Identification of novel transcripts and srna of *Brucella melitensis* by RNA-seq [J]. Chinese Journal of Zoonoses, 2015, 31(3): 216-221.
- [19] 马芳, 刘哲, 康玉军, 等. 虹鳟肝组织新转录本分析及基因结构优化 [J]. 中国实验动物学报, 2019, 27(2): 135-142.
Ma F, Liu Z, Kang Y J, et al. Analysis of novel transcripts and optimization of the gene structure in the liver of rainbow trout [J]. Acta Laboratorium Animalis Scientia Sinica, 2019, 27(2): 135-142.
- [20] Livak K J, Schmittgen T D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C(T)}$ method [J]. Methods, 2001, 25(4): 402-408.
- [21] 兰道亮, 熊显荣, 位艳丽, 等. 基于 RNA-Seq 高通量测序技术的牦牛卵巢转录组研究: 进一步完善牦牛基因结构及挖掘与繁殖相关新基因 [J]. 中国科学(生命科学), 2014, 44(3): 307-317.
Lan D L, Xiong X R, Wei Y L, et al. Yak ovary transcriptome research based on RNA-Seq high-throughput sequencing technology; further improve the yak gene structure and discover new genes related to reproduction [J]. Scientia Sinica (Vita), 2014, 44(3): 307-317.
- [22] 黄兴. 牦牛大脑和小脑低氧适应性的转录组研究 [D]. 成都: 西南民族大学, 2019.
Huang X. The Transcriptome analysis of cerebrum and cerebellum reveals the well hypoxic adaptability of yak [D]. Chengdou: Southwest Minzu University, 2019.
- [23] Yuan Z, Ge L, Sun J, et al. Integrative analysis of Iso-Seq and RNA-seq data reveals transcriptome complexity and differentially expressed transcripts in sheep tail fat [J]. Peer J, 2021, 9: e12454.
- [24] Elsik C G, Tellam R L, Worley K C, et al. The genome sequence of taurine cattle: a window to ruminant biology and evolution [J]. Science, 2009, 324(5926): 522-528.
- [25] 王洪程, 梅楚刚, 咎林森, 等. 牛全基因组测序研究进展 [J]. 西北农林科技大学学报(自然科学版), 2015, 43(11): 17-23.
Wang H C, Mei C G, Zan L S, et al. Progress on whole genome sequencing of bovine [J]. Journal of Northwest A&F University(Natural Science Edition), 2015, 43(11): 17-23.
- [26] 张小强, 沙玉柱, 王正文, 等. 平凉红牛及其杂交和牛的肉品质特征研究 [J]. 中国畜牧杂志, 2021, 57(5): 113-116.
Zhang X Q, Sha Y Z, Wang Z W, et al. Study on the meat quality characteristics of Pingliang red cattle and crossbred Wagyu cattle [J]. Chinese Journal of Animal Science, 2021, 57(5): 113-116.
- [27] Chen J, Guccini I, Di Mitri D, et al. Compartmentalized activities of the pyruvate dehydrogenase complex sustain lipogenesis in prostate cancer [J]. Nat Genet, 2018, 50(2): 219-228.
- [28] 吴非凡, 茆建昱, 丁洛阳, 等. 影响羊肉 pH 变化的因素及其糖原代谢通路机制的研究进展 [J]. 动物营养学报, 2020, 32(2): 571-577.
Wu F F, Mao J Y, Ding L Y, et al. Research progress on influencing factors of mutton pH change and its glycogen metabolism pathway mechanism [J]. Chinese Journal of Animal Nutrition, 2020, 32(2): 571-577.
- [29] Depoortere I. Taste receptors of the gut: emerging roles in health and disease [J]. Gut, 2014, 63(1): 179-190.
- [30] Hong A W, Meng Z, Guan K L. The Hippo pathway in intestinal regeneration and disease [J]. Nat Rev Gastroenterol Hepatol, 2016, 13(6): 324-337.
- [31] Imanaka T. Biogenesis and function of peroxisomes in human disease with a focus on the ABC transporter [J]. Biol Pharm Bull, 2019, 42(5): 649-665.
- [32] Poleti M D, Regitano L, Souza G, et al. Longissimus dorsi muscle label-free quantitative proteomic reveals biological mechanisms associated with intramuscular fat deposition [J]. J Proteomics, 2018, 179: 30-41.
- [33] Wu W, Zhang J, Zhao C, et al. CTRP6 Regulates porcine adipocyte proliferation and differentiation by the adipoR1/MAPK signaling pathway [J]. J Agric Food Chem, 2017, 65(27): 5512-5522.
- [34] Lahalle A, Lacroix M, De Blasio C, et al. The p53 pathway and metabolism: the tree that hides the forest [J]. Cancers (Basel), 2021, 13(1): 133.
- [35] Wu W, Ji M, Xu K, et al. Knockdown of *CTRP6* reduces the deposition of intramuscular and subcutaneous fat in pigs via different signaling pathways [J]. Bba-Mol Cell Biol L, 2020, 1865(8): 158729.