

# 槐猪与杜洛克猪背部皮下脂肪转录组分析

叶鼎承<sup>1</sup>, 杨建敏<sup>2</sup>, 甘乾福<sup>3</sup>, 刘景<sup>1</sup> (1.福建省农业科学院畜牧兽医研究所, 福建福州 350013; 2.福建省福州市长乐区猴屿乡畜牧兽医站, 福建福州 350200; 3.福建农林大学动物科学学院, 福建福州 350002)

**摘要** [目的]了解槐猪和杜洛克猪背部皮下脂肪组织脂肪沉积的分子机制差异。[方法]对槐猪和杜洛克猪背部皮下脂肪组织进行转录组测序, 并进行生物信息学分析。[结果]槐猪和杜洛克猪分别获得 37 858 968 和 37 545 394 条干净读段(clean reads)。槐猪和杜洛克猪分别预测到 9 612 和 9 457 可变剪接事件; 两猪种间共有 709 个差异表达基因, 与杜洛克猪相比, 槐猪有 395 个上调基因和 314 个下调基因。差异表达基因共归类到 46 个 GO 条目中, 其中细胞组分 11 个、分子功能 12 个及生物学过程 23 个; 显著富集的 KEGG 通路共计 9 个, 其中含有与调节脂类代谢显著相关通路, 如胰岛素信号通路和 PPAR 信号通路等。[结论]该研究可为今后研究脂肪沉积的遗传机理提供科学依据。

**关键词** 槐猪; 杜洛克猪; 背部皮下脂肪组织; 差异表达; 转录组测序

**中图分类号** S828; S813 **文献标识码** A

**文章编号** 0517-6611(2020)24-0099-03

**doi:** 10.3969/j.issn.0517-6611.2020.24.027



开放科学(资源服务)标识码(OSID):

## Transcriptome Analysis of Back Subcutaneous Fat of Huai Pig and Duroc Pig

YE Ding-cheng<sup>1</sup>, YANG Jian-min<sup>2</sup>, GAN Qian-fu<sup>3</sup> et al (1. Institute of Animal Husbandry and Veterinary Medicine, Fujian Academy of Agricultural Sciences, Fuzhou, Fujian 350013; 2. Animal Husbandry and Veterinary Station in Houyu Township of Changle District in Fuzhou City, Fuzhou, Fujian 350200; 3. College of Animal Sciences, Fujian Agriculture and Forestry University, Fuzhou, Fujian 350002)

**Abstract** [Objective] To understand the molecular mechanism difference of fat deposition in the subcutaneous adipose tissue between Huai Pig and Duroc Pig. [Method] Transcriptome sequencing of subcutaneous adipose tissue in Huai Pig and Duroc Pig were performed and bioinformatic analysis was made. [Result] 37 858 968 and 37 545 394 clean reads were obtained from Huai Pig and Duroc Pig, respectively. 9 612 and 9 457 alternative splicing events were predicted in Huai Pig and Duroc Pig respectively. There were 709 differentially expressed genes (DEGs) in these two pig breeds. Compared with Duroc Pig, there were 395 up-regulated genes and 314 down-regulated genes in Huai Pig. The DEGs were classified into 46 gene ontology (GO) entries, including 11 cell components, 12 molecular functions and 23 biological processes. A total of 9 pathways had enriched KEGG, which contained pathways significantly correlated with regulation of lipid metabolism, such as insulin signaling pathway and PPAR signaling pathway. [Conclusion] This study can provide scientific basis for the future study of the genetic mechanism of fat deposition.

**Key words** Huai Pig; Duroc Pig; Back subcutaneous adipose tissue; Differential expression; Transcriptome sequencing

脂肪沉积性状是猪主要的经济性状之一, 且与经济价值紧密联系。槐猪(Huai Pig)是福建省特色猪品种, 早熟易肥, 肉质优良, 抗逆性强, 但生长速度慢, 瘦肉率低, 是典型的脂肪型猪种, 而西方现代猪种杜洛克猪(Duroc Pig)生长速度快, 瘦肉率高, 但肉质和抗逆性相对较差, 是典型的瘦肉型猪种。为了解槐猪和杜洛克猪皮下脂肪组织脂肪沉积的分子机制差异, 笔者对这两个猪种背部皮下脂肪进行了转录组测序, 并进行了生物信息学分析。

## 1 材料与与方法

**1.1 试验材料与测序** 选择在相近饲养条件下同期饲养的 190 日龄槐猪和杜洛克猪各 3 头, 屠宰后分别取倒数第 3~4 根肋骨之间的背部皮下脂肪组织, 置于液氮中贮存; 采用 TRIzol 法提取各样品的总 RNA, 并采用 Nanodrop、Agilent 2100 方法对 RNA 样品的纯度、浓度和完整性进行检测, 将符合测序要求的槐猪和杜洛克猪的总 RNA 分别等量混合, 形成 2 个总 RNA 池, 送交北京百迈客生物科技有限公司进行转录组测序(RNA-seq)。

## 1.2 数据整理和分析

**1.2.1 原始数据去杂。** 测序所获得的原始序列数据均带有

一段 3' 接头序列, 并且含有少量低质量序列以及各种杂质成分, 经过数据处理得到可用的干净读段(clean reads)。具体处理方法如下: 去除 3' 接头序列、去除空载读长、去除低质量读段(含有未知碱基 N 的读段)、去除长度过小或过大的读段等<sup>[1]</sup>。

**1.2.2 序列比对和基因表达量分析。** 使用 TopHat 软件将获得的干净读段(clean reads)与猪参考基因组(*Sus scrofa* 10.2; ftp://ftp.ensembl.org/pub/release-75/fasta/sus\_scrofa/)进行比对, 并对比对效率进行统计<sup>[2]</sup>。采用 FPKM (fragments per kilobase of transcript per million fragments mapped) 作为衡量转录本或基因表达水平的指标<sup>[3]</sup>, FPKM 计算公式如下:

$$FPKM = \frac{\text{基因区的读段数目} \times 10^9}{\text{基因长度} \times \text{测序深度}} \quad (1)$$

**1.2.3 可变剪切预测。** 使用 Cufflinks 软件对基因原有的剪接模型进行比较, 进行可变剪接事件(alternative splicing events) 预测<sup>[3]</sup>。常见的可变剪接可分为 6 种类型: 外显子跳跃(exon skipping, ES)、内含子保留(intron retention, IR)、第一个外显子可变剪接(alternative first exon, AFE)、最后一个外显子可变剪接(alternative last exon, ALE)、5' 端可变剪接(alternative 5' splice site, A5SS)、3' 端可变剪接(alternative 3' splice site, A3SS)<sup>[4-5]</sup>。

**1.2.4 基因功能注释及差异表达基因筛选。** 使用 BLASTX 软件将基因序列与 Nr、Swiss-Prot、GO、COG 和 KEGG 数据库

**基金项目** 福建省自然科学基金项目(2012J01111); 福建省公益类科研院所专项(2014R1023-12)。

**作者简介** 叶鼎承(1980—), 男, 福建福州人, 助理研究员, 硕士, 从事畜禽养殖与动物营养研究。

**收稿日期** 2020-08-26

进行比对,获得所有基因的注释信息。使用 R 软件的 EBSeg 包进行基因差异表达分析<sup>[6]</sup>,以  $|\log_2(FC)| \geq 1$  且  $FDR < 0.01$  作为筛选差异基因的标准。其中,差异倍数(fold change, FC)表示两样品之间表达量的比值,错误发现率(false discovery rate, FDR)是对差异显著性  $P$  值进行 Benjamini-Hochberg 方法校正后得到的<sup>[7]</sup>。

**1.2.5 GO 功能分类。**使用 Blast2GO 软件对差异表达基因进行 GO 功能分类,参数  $e$  值  $\leq 10^{-5}$ <sup>[8]</sup>,并使用 WEGO 软件进行统计并绘图。GO (gene ontology) 数据库能对基因、蛋白质进行描述,其包含 3 个本体:细胞组分 (cellular component)、分子功能 (molecular function) 和生物学过程 (biological process)。

**1.2.6 KEGG 通路显著性富集分析。**以 KEGG 通路为单位,找出与整个基因组背景相比在差异表达基因中显著性富集的通路,其计算公式如下:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}} \quad (2)$$

式中,  $N$  为所有基因中具有通路注释的基因数目;  $n$  为具有通路注释基因中差异表达基因的数目;  $M$  为所有基因中注释为某特定通路的基因数目;  $m$  为注释为某特定通路的差异表达基因数目。  $Q \leq 0.05$  的通路定义为在差异表达基因中显著富集的通路 ( $Q$  表示错误比例,  $Q = V/R$ , 其中  $R$  代表所挑选的差异基因个数,  $V$  代表没有差异表达的基因个数, 即假阳性结果)<sup>[1]</sup>。

## 2 结果与分析

**2.1 测序数据质量评估** 对槐猪和杜洛克猪背部皮下脂肪组织进行转录组测序, 获得原始读段数目分别为 46 891 554 和 45 309 436 条, 原始读段经过去杂处理后, 获得的干净读段数分别为 37 858 968 和 37 545 394 条, 其中 78.54% 和 82.08% 的干净读段 (clean reads) 比对上猪参考基因组, 可以满足后续分析的需要。

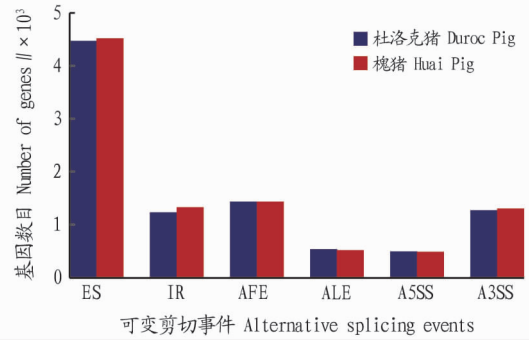
**2.2 可变剪切预测** 槐猪和杜洛克猪背部皮下脂肪组织分别预测到 9 612 和 9 457 可变剪接事件。由图 1 可知, 在 6 个常见的可变剪切类型中外显子跳跃的数量最多。

**2.3 差异表达基因筛选和注释** 对槐猪和杜洛克猪背部皮下脂肪组织进行转录组测序, 并进行差异表达基因筛选。结果发现, 两猪种共有 709 个差异表达基因, 与杜洛克猪相比槐猪有 395 个上调基因和 314 个下调基因。各个数据库注释到的差异表达基因数目及所占比例如表 1 所示。

**2.4 GO 功能分类** 由图 2 可知, 槐猪和杜洛克猪背部皮下脂肪组织的差异表达基因共归类到 46 个 GO 条目中, 其中细胞组分 11 个、分子功能 12 个及生物学过程 23 个。细胞组分中分类到较多基因的条目为 cell part、cell 和 organelle, 分别含 540、540 和 394 个基因; 分子功能中的条目 binding 含 450 个基因; 生物学过程中的条目 cellular process、metabolic process 和 biological regulation 分别含 519、462 和 412 个基因。

**2.5 KEGG 通路显著性富集分析** 槐猪和杜洛克猪背部皮

下脂肪组织之间的差异表达基因共显著富集到 9 个 KEGG 通路, 通路信息如表 2 所示。



注: ES. 外显子跳跃; IR. 内含子保留; AFE. 第一个外显子可变剪接; ALE. 最后一个外显子可变剪接; A5SS. 5' 端可变剪接; A3SS. 3' 端可变剪接

Note: ES. Exon skipping; IR. Intron retention; AFE. Alternative splicing of the first exon; ALE. Alternative splicing of the last exon; A5SS. Alternative splicing of the 5' end; A3SS. Alternative splicing of the 3' end

图 1 可变剪切事件统计

Fig.1 The statistics of alternative splicing events

表 1 差异表达基因注释统计

Table 1 The annotation statistics of DEGs

数据库 Database	差异表达基因数目 Number of differentially expressed genes	占比 Proportion // %
Nr	691	97.46
Swiss-Prot	627	88.43
GO	642	90.55
KEGG	402	56.70
COG	264	37.24

## 3 讨论

该研究对槐猪和杜洛克猪背部皮下脂肪组织进行转录组测序, 分别获得了 37 858 968 和 37 545 394 条干净读段。陈伟<sup>[9]</sup>对莱芜猪和大白猪背最长肌的转录组测序分别获得 40 498 476 和 59 072 892 条干净读段。Sodhi 等<sup>[10]</sup>对济州岛本地猪和巴克夏猪的脂肪组织进行了测序, 分别获得 41 411 459 和 40 308 452 条干净读段。由此可见, 对不同猪品种同一个组织测序所获得的干净读段数目不同; 对同一个猪品种的不同组织的测序所获得的干净读段数目也不同, 这可能是由于不同组织所行使的生理功能和所含的基因数目不同。

该研究在槐猪和杜洛克猪 2 个品种猪背部皮下脂肪中分别预测到 9 612 和 9 457 个可变剪切事件, 其中外显子跳跃的数量最多。Kang 等<sup>[11]</sup>在研究不同年龄滩羊皮肤组织的可变剪切时, 在 1 月龄皮肤组织发现有 4 120 个可变剪切, 而在 48 月龄的皮肤组织中有 4 237 个可变剪切事件。李青芝等<sup>[12]</sup>在长白母猪背部皮下脂肪、大网膜脂肪和肠系膜脂肪的转录组研究中发现 936~1 029 个可变剪切事件。对基因组中可变剪切事件的预测, 将有助于理解这些基因在转录调控机制方面的功能。

该研究对槐猪和杜洛克猪背部皮下脂肪组织之间的差

异表达基因进行 GO 功能分类和 KEGG 通路分析,有助于进一步了解差异表达基因的功能及基因产物在细胞代谢过程中的作用。KEGG 通路反映蛋白质之间的互相作用关系,参与代谢通路的蛋白的表达量或者酶的活性改变都可以引起通路调节出现差异。该研究显著富集的通路中,发现一些与

调节脂类代谢显著相关通路,如胰岛素信号通路和 PPAR 信号通路等,为进一步研究脂肪沉积的分子机制提供了一定的参考,后期将结合蛋白质功能试验(如基因敲除、转基因和 RNA 干扰等)进一步证实。

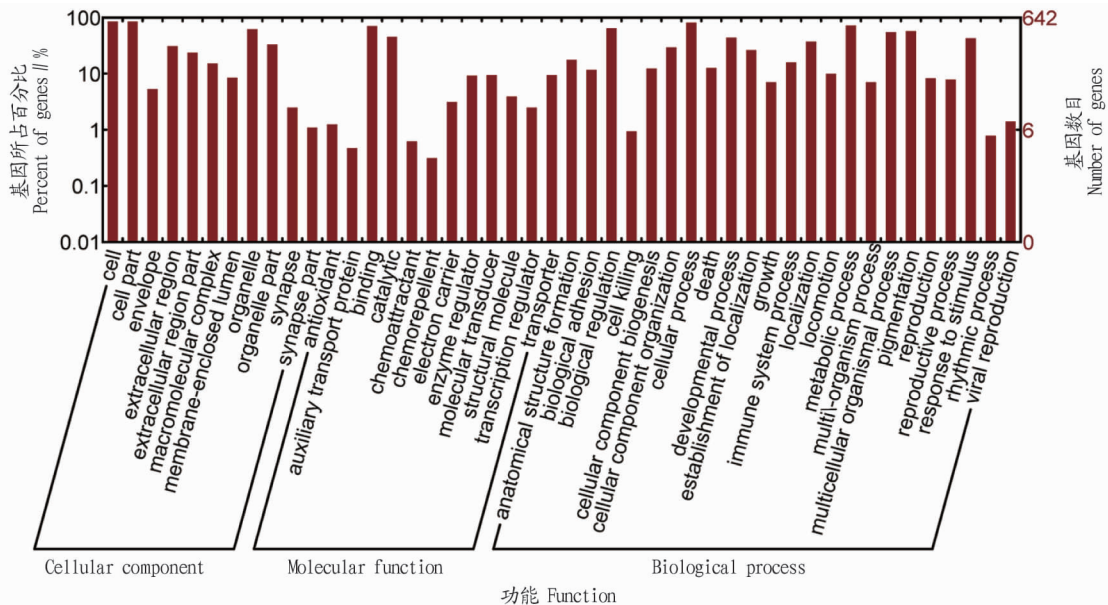


图2 差异表达基因 GO 功能分类

Fig.2 GO function classification of DEGs

表2 差异表达基因显著富集的9个通路

Table 2 Enriched 9 pathways of DEGs

通路名称 Pathway name	通路中有注释的 差异表达基因数目 Number of DEGs annotated in the pathway	Q 值 Q value	通路 ID 号 ID number of pathways
补体系统 Complement and coagulation cascades	32	0.00E+00	ko04610
药物代谢-细胞色素 P450 Drug metabolism-cytochrome P450	21	3.40E-11	ko00982
视黄醇代谢 Retinol metabolism	20	1.39E-10	ko00830
异生素细胞色素 P450 新陈代谢通路 Metabolism of xenobiotics by cytochrome P450	17	5.88E-08	ko00980
PPAR 信号通路 PPAR signaling pathway	12	8.10E-05	ko03320
甾类激素的生物合成 Steroid hormone biosynthesis	22	2.38E-04	ko00140
胰岛素信号通路 Insulin signaling pathway	11	3.02E-04	ko04910
淀粉和蔗糖代谢 Starch and sucrose metabolism	13	4.33E-04	ko00500
糖酵解/糖异生 Glycolysis / Gluconeogenesis	12	2.14E-03	ko00010

## 参考文献

- [1] 耿立英, 巩元芳, 刘铮铸, 等. 北京油鸡和来航鸡脾脏差异表达基因的筛选与分析[J]. 中国兽医学报, 2017, 37(5): 941-948, 964.
- [2] KIM D, PERTEA G, TRAPNELL C, et al. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions[J]. Genome Biol, 2013, 14(4): 1-13.
- [3] TRAPNELL C, WILLIAMS B A, PERTEA G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation[J]. Nat Biotechnol, 2010, 28(5): 511-515.
- [4] 费虹虹. 纹枯病菌对寄主侵染的选择性机制初步研究[D]. 雅安: 四川农业大学, 2015.
- [5] 刘文鑫. 杨树溃疡病发生的水分生理效应及转录组分析[D]. 保定: 河北农业大学, 2015.
- [6] LENG N, DAWSON J A, THOMSON J A, et al. EBSeg: An empirical Bayes hierarchical model for inference in RNA-seq experiments[J]. Bioinformatics, 2013, 29(8): 1035-1043.
- [7] 李国辉. 金茅黑鸡 A 系胸肌转录组特征及对重要肉用性状的影响[D]. 扬州: 扬州大学, 2018.
- [8] GÖTZ S, GARCÍA-GÓMEZ J M, TEROL J, et al. High-throughput functional annotation and data mining with the Blast2GO suite[J]. Nucleic Acids Res, 2008, 36(10): 3420-3435.
- [9] 陈伟. 莱芜猪和大白猪背最长肌 miRNA 与 mRNA 转录组测序及特征分析[D]. 泰安: 山东农业大学, 2014.
- [10] SODHI S S, PARK W C, GHOSH M, et al. Comparative transcriptomic analysis to identify differentially expressed genes in fat tissue of adult Berkshire and Jeju Native Pig using RNA-seq[J]. Mol Biol Rep, 2014, 41(9): 6305-6315.
- [11] KANG X L, LIU G, LIU Y F, et al. Transcriptome profile at different physiological stages reveals potential mode for curly fleece in Chinese tan sheep[J]. PLoS One, 2013, 8(8): 1-9.
- [12] 李青芝, 王滔, 李明洲, 等. 猪皮下与内脏脂肪组织 mRNA 转录组的构建与差异分析[J]. 四川农业大学学报, 2013, 31(1): 65-70.