

申 报	系列：实验技术
	专业：畜牧学
	职称：正高级实验师

业绩成果材料

（申报人的业绩成果材料包括论文、科研项目、获奖以及其他成果等）

单 位
(二级单位)

动物科学学院
(国家生猪种业工程技术研究中心)

姓 名

郑恩琴

材料核对人：

单位盖章：

核对时间：

华南农业大学制

目 录

一、教学研究业绩

1. 教改论文：基于微课的猪生产学实验课程混合教学研究
(检索证明) 1

二、科研项目

1. 主持：国家农业农村部重大科技计划项目“体型性状相关功能基因与分子标记筛选、鉴别及应用”子课题合同书.....9
2. 主持：国家重点领域研发计划项目“高繁大黑猪新品种培育及繁育体系构建”子课题合作协议.....23
3. 主持：广东省重点领域研发计划项目“种猪杂交育种技术创新与高繁优质新品系培育”之“杂交育种与现代分子设计育种技术研发”课题任务书..... 30
4. 主持：广东省自然科学基金面上项目“猪 12 号染色体影响初生重主效 QTL 的分子遗传机理解析”项目任务书.....34
5. 主持：畜禽育种国家重点实验室开放课题“游离氨基酸对猪肉风味影响及其在肉品内的形成机制”合同书.....42
6. 主持：广东省科学技术厅“百千万工程”农村科技特派员项目合同书.....47
7. 主参：国家自然科学基金面上项目“鉴别猪 13 号染色体影响悬蹄过度生长的目的基因及其关键突变”59
8. 主参：国家自然科学基金青年科学基金项目“组蛋白 H3K27me3 对猪胚胎骨骼肌发育调控作用的研究”67

三、论文

1. 检索证明.....	74
2. 以通讯作者发表本专业论文情况	
2.1. Protein Dynamic Landscape during Mouse Mammary Gland Development from Virgin to Pregnant, Lactating, and Involuting Stages , Journal of Agricultural and Food Chemistry. 2024, 72, 7546–7557.....	79
2.2. Genome-wide detection of multiple variants associated with teat number in French Yorkshire pigs , BMC Genomics. 2024, 25:722.....	91
2.3. Increased Accuracy of Genomic Prediction Using Preselected SNPs from GWAS with Imputed Whole-Genome Sequence Data in Pigs, Animals 2023, 13, 3871.....	103
2.4. Identification of Candidate Genes for Economically Important Carcass Cutting in Commercial Pigs through GWAS, Animals, 2023, 13, 3243. doi.org10.3390ani 13203243.....	115
2.5. Identification of candidate genes associated with carcass component weights in commercial crossbred pigs through a combined GWAS approach , Journal of Animal Science, 2023, 101, 1–12.....	128
2.6. Integrated Single-Trait and Multi-Trait GWASs Reveal the Genetic Architecture of Internal Organ Weight in Pigs ,	

Animals 2023, 13, 808. doi.org10.3390/ani13050808.....	140
2.7.Treatment of Donor Cells with Oxidative Phosphorylation Inhibitor CPI Enhances Porcine Cloned Embryo Development, Animals. 2024, 14(9), 1362. doi.org/10.3390/ani14091362.....	153
2.8. Genome-Wide Association Studies for Flesh Color and Intramuscular Fat in (Duroc X Landrace X Large White) Crossbred Commercial Pigs, Genes, 2022, 13, 2131. doi.org10. 3390 genes13112131.....	165
2.9. 染色质转座酶可及性测序研究进展, 遗传, 2020, 042 (004):333-346.....	175
2.10. 瘦肉型猪基因组中 ROH 的应用及研究进展, 中国畜牧 杂志, 2022, 58.10:1-7.....	192
2.11. 不同饲养模式对母猪分娩前后行为影响的研究, 家畜 生态学报, 2019,40(5)40-44	205
2.12. Gender Control of Mouse Embryos by Activation of TLR7/8 on X Sperm via Ligands dsRNA-40 and dsRNA-DR, Molecules, 2024, 29, 262. doi.org/10.3390/molecules29010262	215
2.13. A Transcriptome Analysis Reveals that Hepatic Glycolysis and Lipid Synthesis Are Negatively Associated with Feed Efficiency in DLY Pigs, Scientific Reports (2020) 109874,doi. org10. 1038s41598-020-66988-6.....	228

2.14. Brain Transcriptome Analysis Reveals Potential Transcription Factors and Biological Pathways Associated with Feed Efficiency in Commercial DLY Pigs, Dna and Cell Biology, 2021, 40(2), DOI 10.1089dna. 2020.6071.....	240
--	-----

四、科研成果

1. 科技奖励证书

1. 1. 广东省科技进步奖一等奖：瘦肉型种猪基因组育种技术与应用.....	251
1. 2. 高等学校科学研究优秀成果奖（科学技术）一等奖：新型高效“温氏 WS501 猪配套系”培育与应用.....	252
1. 3. 神农中华农业科技奖一等奖：五系配套瘦肉型猪选育关键技术与应用.....	253
1. 4. 广东省农业技术推广奖一等奖：五元杂交瘦肉型种猪新配套系及其养殖技术示范推广.....	254

2. 知识产权

2. 1. 专利授权证书：一种用于肉质样品测定预处理的恒温水浴装置.....	255
2. 2. 专利授权证书：一种猪 1 号染色体上与猪剩余采食量相关的 SNP 分子标记及其用途.....	256
2. 3. 专利授权证书：一种影响猪饲料转化效率性状的 SNP 分子标记及其用途.....	257
2. 4. 专利授权证书：一种猪 11 号染色体上影响猪背膘厚的	

拷贝数变异分子标记及应用.....	258
2.5. 专利授权证书：一种与猪剩余采食量相关的 SNP 分子标记及其用途.....	259
2.6. 专利授权证书：一种位于猪 14 号染色体上与母猪死胎数和健仔率相关的 SNP 分子标记及其用途.....	260
2.7. 专利授权证书：一种猪 13 号染色体上与母猪产畸形仔猪数相关的 SNP 分子标记及其用途.....	261
2.8. 专利授权证书：一种猪 7 号染色体上与猪死胎数、活仔率相关的 SNP 分子标记及其用途.....	262
2.9. 专利授权证书：一种猪 3 号染色体上与猪日增重和上市体重日龄相关的拷贝数变异分子标记及应用.....	263
2.10. 专利授权证书：位于猪 16 号染色体上与猪瘦肉率和眼肌面积相关的 SNP 分子标记及应用.....	264
2.11. 专利授权证书：位于猪 7 号染色体上与总乳头数相关的 SNP 分子标记及应用.....	265
2.12. 专利授权证书：猪 6 号染色体上与瘦肉率、眼肌面积、眼肌厚度相关的 SNP 分子标记与应用.....	266
2.13. 专利授权证书：位于猪 7 号染色体上与杜洛克猪日增重性状相关的分子标记及应用.....	267

五、其他业绩

1. 华南农业大学动物科学学院 2020 年青年教师教学观摩比赛一等奖.....	268
--	-----

主管 | 湖北省科学技术协会

主办 | 湖北科教导刊杂志社

国内统一刊号: CN42-9001/N

国际标准刊号: ISSN 1674-6813

科 教 导 刊

电
子
版

34期
2019/¹²_{上旬刊}

科教导刊 (电子版)

The Guide of Science & Education (Electronic Edition)

目次

2019 年第 34 期
12 月上旬刊

主 管：湖北省科学技术协会
主 办：湖北科教导刊杂志社
编辑出版：湖北科教导刊杂志社

总 编：馥品
编 辑 部：李政 朱敬 古小燕 殷汉思
美术编辑：易晓敏

地 址：武汉市洪山区珞狮北路 76 号
书香门第大厦 1205 室
邮政编码：430070
电 话：027-87508018 027-87508618
电子信箱：kjdkdzb@126.com
网 址：dzb.kjdlkzss.com

发行单位：武汉中淼文化传媒有限公司
定 价：RMB16.00 元
印 刷：武汉贝思印务设计有限公司
国内统一刊号：CN 42-9001/N
国际标准刊号：ISSN 1674-6813

* 万方数据—数字化期刊群收录期刊
* 中国核心期刊(遴选)数据库收录期刊
* 龙源期刊网收录期刊
* 《中文科技期刊数据库》收录期刊

本刊公告

为适应我国信息化建设，扩大本刊及作者知识信息交流渠道，本刊已被多个数据库收录。来稿凡经录用，如无特殊声明，即视作投稿者同意授权本刊及本刊合作媒体进行信息网络传播及发行。

◆ 科教前沿

“人工智能+教育”时期高校学科建设新探	贾超广 肖海霞 001
科技创新资源对大学生自主创新能力培养的价值	李志勇 002

◆ 教育观点

在校大学生学习内驱力影响因素研究 ——以杭州市 5 所高校为例	杜 磊 翁异静 003
新媒体时代对高校思政课立德树人的挑战与应对策略	杜伟伟 006
孔子“绝四”与大学生和谐关系的构建	周家乐 张 岩 008
大学生“宅”现象调查研究 ——基于重庆 C 高校的问卷调查	向丽虹 石嘉欣 周代玖 010
高校思政课开展生态文明教育的对策研究	高 迪 012
高职院校“思政课程”与“课程思政”的协调机制构建研究	余 璐 徐燕秋 014
学生社团与学习共同体切合点及促进学习效果检测研究	胡彦军 尚 峥 孙玉强 016
新媒体背景下高校学生干部培养探究	刘海燕 018
新时期高职院校建设发展的“瓶颈”问题研究与对策	汤永刚 020
新时期高校智慧教室建设与教育应用分析	李 龙 李小强 刘国瑞 延 清 021
新理念下中职感恩教育开展的思路分析	王宛珊 022
华德福教育中环境创设思想探析	吴玲玲 023
单亲家庭大学生的心理现状及对策研究	周沫含 025
依托湘江新区开拓地方高校创新创业人才培养模式探讨	陈明辉 026
幼儿建构活动中“留白”策略运用初探	钟晓琦 028
课堂提问要注意到学生个性差异	隋 荣 029
家庭教育中回应 3-6 岁幼儿哭泣行为的研究	吴炳贞 030
知识与权力的互动 ——《教育与社会分层》读后感	戚媛媛 031
基于在线旅游企业的高职旅游人才探析	陈 妍 冯淑玲 032
团体沙盘游戏疗法在中学生心理健康教育中的展望	麻伟婷 034
浅谈小学一年级新生“一口常规”的规范	宗凡滢 035
浅析“儒厨”培养理念下的人才培养机制	李洪磊 周建龙 036
高职院校新生心理健康测试后的追踪研究	李 斌 037
“互联网+”下学生创新创业能力培养研究	段玮林 038
三亚市蓝鲸幼儿园科学实验课存在的问题	曹惠容 王 兵 魏青云 贺 婕 039
探究“校企合作、产教融合”一体化创新人才培养新模式	崔晓丽 崔红卫 潘尧坤 040
校企合作-以就业为导向培养技能型人才	王韬旭 042
高校实施政府会计制度的难点与建议	郭同燕 043

基于大数据技术精准定位人才培养规格的研究与实践
——职业能力分析大数据服务平台应用 周士凯 044

高校大学生创新创业能力培养的策略研究
——以喀什大学为例 彭 慧 045

“互联网+”背景下,地方高校开展大学生创新创业教育的现状分析 苏明磊 046

基于企校合作的德国“双元制”农业职业教育的审视方法初探 郑拥民 047

新工科下本科物联网专业培养模式研究 卢 岩 048

互联网负面信息对青少年危害及其整治的必要性研究 杨 宏 049

试析大学传统教学中的弊端及解决策略 贺碧圆 050

公共艺术教育与专业艺术教学紧密结合的创新实践探索 包文亚 051

对高校学生心理健康问题的一点思考 曹庆琼 052

中医药教育视角下大学生非物质文化遗产传承认知研究 李 毅 冯 春 朱斯斯 罗欣延 053

“新商科”理念下高校电子商务人才的培养 梁海跃 055

中外教师课堂鼓励性用语比较 张雅琼 056

专业硕士实践教学困境分析及出路探索 蔡 强 057

小学生劳动技能形成之我见 郑云成 058

东盟国家高层次来华留学生教育发展分析 邹德虎 059

论地方课程资源的基本特性 董乾然 060

新时代下做好疾控系统青年职工思想政治教育关键环节探析 李 思 胡 伟 戴凤屏 王德明 061

◆ 教育管理

非洲来华留学生教育管理工作研究 何 洁 062

论技工院校学籍档案的信息化管理 余奕松 064

青海高校构建辅导员多元文化素质的必要性 宋 云 066

彰显家国情怀的“半军事化”高职学生管理体系构建与实践 张 伟 温长胜 068

质的研究对促进教师教育教学的重要性 张建东 070

刍议幼儿教师提升音乐素养的策略 张长杰 越小顶 071

大数据对我国高校教育管理的影响及对策研究 鲍松堂 072

以爱润心,以德导行
——浅析小学班主任教育工作中的心理健康教育 欧阳锦弟 073

高中物理实验的实验室安全管理分析 庞永昌 075

“三个教育”提升师德水平 顾艳萍 076

温润童心,插柳成荫
——浅谈新时代下小学班主任的教学管理策略途径 许景荣 077

以学生为主体,推进小学班级管理工作 余 江 079

幼儿教师进行音乐特长训练的必要性研究 越小顶 张长杰 080

浅谈高中德育工作中以德治班的实践方案 袁传能 081

现代远程教育教学管理信息系统的探索与实践探究 杨日新 082

大数据背景下高职院校学生管理工作信息化的运用探讨 丁 娜 083

浅谈高校实验室安全文化建设策略 郑 磊 张 丹 钱婷婷 高 珂 084

新工科背景下高校基础力学教师培养途径探讨 杜春志 刘兵飞 086

浅谈小学班主任德育工作的策略 李美芳 087

高校辅导员做好大学生就业心理教育的思考 刘 乐 088

徽州文会视角下班级凝聚力培育对策研究 夏 莹 090

◆ 课程教学

基于数字图像处理课程内容的实践案例实施
——以“自然环境下的苹果图像的分割与定位”为例 马慧敏 袁 涛 焦 俊 乔 焰 朱 诚 091

在高中化学教学应用微课的成效分析
——以电解池应用的工业生产为例 庞达藕 094

国际经济与贸易专业课程教学内容整合优化的探索与实践
——以浙江农林大学为例 邵来安 096

高校艺术类专业第二课堂课程建设实践思考
——以 Z 大学美术学院为例 李高举 098

综合性大学《曲式与作品分析》课程教学窥探 邹红云 100

勿忘传统,方得始终论传统文化在小学美术教学中的渗透研究
——以顺德传统文化纸通仔为例 黎梓颖 102

基于探究式教学和课外科技活动背景下的“生物药剂学与药物动力学”课程改革初探
..... 葛建林 芳 邓同乐 张永勇 周益峰 李敏杰 104

家乡文化融入高校思政课教学的效能 李佳先 106

谈初中物理课堂中的情境教学 赵权福 108

小学低年级道德与法治教学情境的创设方法 尔孜古丽·毕代 109

大数据时代下的高校思政教学改革 曹宁 110

新工科环境下机械设计制造及其自动化专业建设研究 符巨博 蔡高参 孟培培 杨金林 袁永锋 郭绍义 111

基于“三创平台”的自动化类实训课程教学模式探索与实践 朱铝芬 陈美玲 张云 石瑶 112

小学科学实验材料准备的有效途径 赵世金 113

如何提高小学音乐教学的有效性 李芙蓉 114

宝宝运动在亲子教学中的实践研究 高晶萍 115

怎样通过作业设计来培养中学生的物理核心素养 郑怀 116

基于 STEAM 教育理念的初中信息技术教学分析 邓刚 117

回归生活,学以致用
——探析初中化学生活化教学 张晓霞 118

基于微课的猪生产学实验课程混合教学研究 郑恩琴 洪林君 119

初中信息技术课堂有效教学的思考 李新光 121

提高《生物工程设备》课程教学质量的探索与实践 陈玄阳 郑立霖 李慧星 于海彦 罗建成 122

非电工类专业电工学课程实践教学体系改革探索 徐宁 杨东东 123

小学美术教学中如何培养学生的核心素养 张玉芹 124

高中物理教师实验教具开发现状调查及对策 王磊 125

浅谈小学美术教育中的色彩教学 刘雨婷 126

认知失调理论在高校文科基础理论课教学中的应用 傅立光 127

中华优秀传统文化融入高职学前教育专业课程策略研究 江盼 128

生物工程设备课程建设的思考与实践 高家乐 范玉璞 李慧星 罗建成 129

针对高职计算机教研析创新思维的培养 李燕萍 130

《串联与并联》教学设计 王松飞 131

刚体基本规律的理解和提高 曹云玖 132

新旧高中地理人教版必修一教材比较研究 景晓菲 133

图像法在高一物理必修一中的应用 刘忠旭 134

关于《沪剧之美》主题式 PBL 项目化学习方案设计的实践研究 程玥 136

基于信息时代的桥式整流电路教学设计 夏于琴 137

初中生物教学中趣味教学探究 孙巧绒 138

高校课堂教学初探 罗菊珍 139

《汽车机械基础》平面四杆机构的教学改革研究 廖梅花 140

基于发展核心素养的初中化学有效教学实践研究 姜雪梅 张颖 141

以就业为导向的高职计算机教学模式优化方法浅谈 刘建伟 143

论互联网时代的视唱练耳教学改革 程娟 144

基于创客教育的计算机专业课程校本化实施中实例教学法的应用 王燕 145

职业院校实施专业课“课程思政”的思考 姚岚 叶琦 146

基于实践能力提升与多维互动的市场营销学 SPOC 教学改革 毛凌琳 148

怎样在有限的时间内备课 李芳 149

浅谈关于文科生的医用化学教学 李珊珊 150

“互联网+”环境下高校计算机教学方法探究 罗盛章 151

基于微课的猪生产学实验课程混合教学研究

郑恩琴 洪林君

(华南农业大学动物科学学院/国家生猪种业工程技术研究中心 广东·广州 510642)

摘 要 微课形象直观、资源多样、情境真实、使用方便,能够显著提升学生的学习积极性。把微课与传统实验教学相结合的混合教学能够丰富课程教学形式,提升学生的学习兴趣,拓宽学生的视野,加深学生对课程内容的掌握程度。对提高教学质量具有重要意义。

关键词 微课 猪生产学 实验课 混合教学

中图分类号:G642

文献标识码:A

实验课是提升大学生动手能力和综合专业素质的重要课程,尤其是对于与生产实际紧密联系的农林高校。随着新一代信息技术的迅猛发展,人工智能的时代已经来临,且正深刻影响着农业各领域的发展,这对农林高校在学生培养提出了更高要求,不仅要加深理论知识学习,而且更注重生产实践能力的提升。因此,为适应新农科发展的需要,很多高校都把实验课摆在了更高的地位,猪生产学实验课已是众多农林高校的独立必修课程。随着网络和智能手机的普及,教育信息化迅猛发展,以微课为代表的一批线上教育平台迅速发展。猪生产学实验课不仅要学生掌握理论基础知识,更要求学生掌握实践操作。但是随着目前非洲猪瘟在我国的持续蔓延,几乎所有的养殖场都不接受学生进场实习操作,因此很多实验环节无法在现场开展。微课是以视频教学为主要形式,以线上移动学习的方式,针对课程中某个教学环节或知识点而开展教与学活动的各种资源的有机组合。采用微课的在线学习与传统的面对面教学相融合的混合教学模式,既可以把生产实践中相关的操作技术通过微视频融入到课程教学中来,也能充分激发学生的学习积极性,使学生更好的掌握猪生产学实验课程中的相关技术。因此,微课与传统教学相结合的猪生产学实验课程混合式教学改革具有重要意义,并且可以给农业院校相关的其他实验课的教学改革提供参考。

1 基于微课的猪生产学实验课教学改革的必要性

自2018年8月份在我国境内首次发现非洲猪瘟病毒以来,我国大部分地方都先后有非洲猪瘟疑似病例的报道,因此目前国内绝大部分养猪场都加强了防疫级别,杜绝场外人员进入,也不允许农林高校动物科学专业本科生进场进行生产实践环节的操作学习,而且现在活猪运输管控严格,实验课很难用活猪进行,只能购买白条猪供学生进行实验,从而导致很多生产环节以及猪只屠宰流程等学生都无法知道。因此,在此背景下猪生产学实验课教学急需改革。

微课是根据课程要求以知识点为单位,时长为10分钟左右的视频教学,具有短小高效、资源多样、形象直观、情境真实、指向明确、使用方便等特点,对激发学生学习乐趣,提高自主学习能力很有帮助。微课的优势主要有以下三个方面:(1)突出主题和特定内容。微课虽然教学时间短,但其主题突出,内容清晰、准确。微课内容都是经过教师课前挑选并精心制作而成,能够更好的辅助教师完成课堂教学任务,并且激发学生

的学习积极性,从而更高效的开展课堂教学。(2)教学时间较短。短小高效是微课最突出的特征之一,现在每个人的生活节奏都很快,微课可以很好地适应社会趋势,普遍受到学习者的欢迎。微课视频一般持续6到9分钟,最长也是10分钟左右。学生可以利用空闲时间进行学习,比如食堂就餐间隙,课间休息时间,甚至坐公交地铁的时间都可以学习,从而显著提高了学习效率。(3)教师充分发挥其创造力,增加了学生学习的乐趣。微课具有一定的主动性和创造力,不会受制于传统教育理念的限制和影响,这些因素对于实现良好的教学目标具有重要作用。在猪生产学实验课程教学中使用微课可以极大地提高学生学习的积极性,利用形象有趣的教学视频增强学生的学习主动性,使他们能够成为自己学习的主人。但是微课是纯在线的教学模式,对于实践性要求较低的公共课很适合,但对于猪生产学实验课程这种操作性和实践性很强的专业课,进行纯在线教学达不到教学要求,需要与师生面对面课堂教学结合起来才能起到好的教学效果。因此基于微课的混合式教学更适用猪生产学实验课程的教学改革。

2 基于微课的混合式教学设计

基于微课的混合式教学过程主要包括三个环节:(1)根据课程教学目标和计划,制定相应的学习计划和内容;(2)根据实施混合式教学的设施设备,制定合适的学习内容;(3)实施混合式教学并对教学结果进行评价。根据上述理论,笔者提出基于微课的猪生产学实验课程混合式教学设计流程(如图1所示),按照这个流程,混合式教学过程包括课前在线学习、面对面课堂教学和课后巩固与反馈三个阶段。

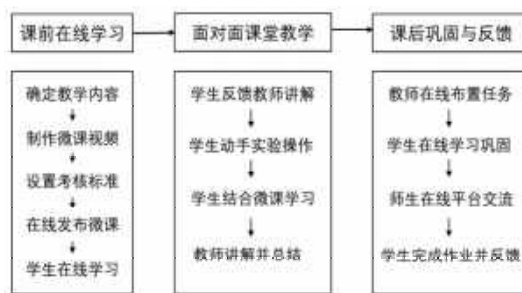


图1：猪生产学实验课程混合式教学流程图

2.1 课前在线学习

根据猪生产学课程实验教学目标,分解教学内容,细化知识点,针对每个知识点请猪场相关人员利用相机录制生产实

践相关视频,如猪活体背膘测定,猪体长、体高、体重、管围等指标的测定。教师也可以从网上获取相关视频。获取原始视频后,教师利用视频处理软件对视频进行后期制作,包括剪接、添加字幕等。然后把制作好的视频与其对应的内容相结合,并在网上发布,并根据教学目标,在在线平台中设置相应的练习题。另外教师也可把本节实验课的难点、重点以及实验操作过程中的注意事项等内容同时在平台中发布,方便学生在实验课前预习,便于充分了解本节实验课的内容以及需重点关注的知识点。这样正式开展课堂面对面教学时就不至于使学生由于对实验内容不熟悉,或理论知识不了解而不敢动手操作实验,或不规范的操作实验。

2.2 面对面课堂教学

学生课前利用手机等设备在线观看教师发布的微课内容,对本次实验课程的内容已经比较熟悉,因此,在课堂上对老师布置的实验操作任务不陌生不惧怕,这样学生动手操作就顺畅多了。实验操作过程中有遗忘的知识点,学生也可以随时观察手机等移动设备进行观看学习,从而提高了实验的准确性和效率,学生也可从实验操作过程中提升自己的动手能力并获得实验成功的喜悦感,进而加深对实验课程的掌握程度。另外,相关生产实践过程如猪活体背膘测定,猪体长、体高、体重等指标的测定以及猪只屠宰流程学生都可以通过微课进行

学习,并在课堂上与老师进行讨论。教师在课堂主要进行针对性的讲解重点、难点,并对学生在学习微课内容反馈的问题进行解答。实验课程结束后,老师对本节课进行总结。

2.3 课后巩固、拓展

教师利用微课发布相关视频资料并布置课后作业,学生在线完成作业并观看相关视频对知识进行巩固拓展。

3 小结

基于微课的猪生产学实验课混合式教学能够把教师精心准备的视频供学生自由学习和讨论,促进了实验课的学习效果,并且能把养猪场相关实践环节带到课堂,在目前学生无法进猪场接触生产环节的情况下最大限度的提升学生对猪只的认识。这种新的教学方式使得学生能够根据自己的学习生活时间来安排学习,给学生带来了很好的学习体验,对提高实验课程教学质量具有重要意义。

参考文献

- [1] 黄虹,魏宸,邹长伟.“微课-慕课-翻转课堂”立体教学模式的构建——以“环境评价”课程为例[J].现代教育技术,2016,26(10):86-92.
- [2] 鲁晓帆.基于微课的计算机网络课程教学设计研究[J].通讯世界,2018,335(04):275-276.
- [3] 王妍莉,马明辉,严瑾.基于 Blackboard 平台的民族高校混合式教学行动研究[J].电化教育研究,2015,36(09):77-82.

(上接第 113 页)习,通过亲身体验,提高学生对科学原理等相关知识的理解,加强学生对科学知识的记忆,这样一来,还能够有效减轻教师搜集实验材料的工作负担,使学生养成细观察、勤思考的习惯,让学生能够全身心的探究科学实验。例如,在科学课程中让学生认识树叶之间的不同,学生通过收集树叶得以了解树木品种,再例如在有关动物的科学课程中,学生可以通过了解不同动物来了解其生活环境、生长特性等,充分发挥了学生的自主能动性,使学生得以主动的对科学知识进行探索研究。

2.4 共同探讨,加强互动

在科学实验教学活动中,教师向学生开展教学活动时,如果直接按照教材内容对学生开展实验活动,可能会导致学生无法有效了解实验原理,不利于提高科学课程的教学效果。在加上实验器材的不足,极易导致实验活动难以继续,不利于提高学生科学素养,针对此现象,教师在教学活动中要多与学生进行沟通,加强与学生的互动,通过替换实验材料的方式,促使实验活动有序开展。例如,教师在对讲讲述有关摆的研究时,让学生思考如果确保摆场不变,改变摆的路径,引导学生采用小桶作为摆的容器,在容器中放入不同重量的物质,这样替换实验材料的方式,能够使学生亲身参与到实验活动中,提高了学生对实验器材的掌握程度。

除了上述方法之外,准备小学科学实验材料的有效途径

远远不止这些,还需要我们不停的探索,小学科学课程的教师在对讲开展实验教学时,要有意识拓宽实验材料的准备途径,为实验活动的顺利开展提供保障,其次,还要在试验过程中引导学生进行课程准备、讨论,加强学生对科学知识的记忆,以便培养学生创造性思维,不断提高学生的科学素养。

3 结语

综上所述,教师在对小学生开展科学实验教学活动时,为了提高实验教学效果,教师要有效准备各项实验材料,通过实验活动加强学生对科学知识的理解与记忆,此外,使学生亲身参与到实验活动中也能够激发学生的学习积极性,教师在于学生互动式,有意识的引导学生掌握科学知识,不断提高学生的科学素养,为学生的学习生涯提供保障。

参考文献

- [1] 李权.选用合适材料提高小学科学实验学习的有效性[J].小学科学(教师版),2018(03):26.
- [2] 刘其容.浅探小学科学课实验材料准备的方法[J].小学科学(教师版),2015(07):3.
- [3] 李加炎,王彦.怎样做好小学科学实验器材与素材的教学准备[J].实验教学与仪器,2019,36(04):63-64.
- [4] 宗全.小学科学有效实验教学建构的准备策略[J].教育与装备研究,2015(09):32-34.
- [5] 廖传忠.小学科学实验活动中学生观察能力培养的策略[J].新课程(小学),2015(01):195.

电子|版

科 教 导 刊

主 管：湖北省科学技术协会

主 办：湖北科教导刊杂志社

出 版：湖北科教导刊杂志社

总 编：馥品

编 辑 部：李政 朱敬 古小燕 殷汉思

美术编辑：易晓敏

刊 期：旬刊

地 址：武汉市洪山区珞狮北路76号书香门第大厦1205室

电 话：027-87508018 027-87508618

投稿邮箱：kjkdzb@126.com

印 刷：武汉贝思印务设计有限公司

发 行：武汉广洲广告有限公司

ISSN 1674-6813



SCAULIB202405576

检索证明

根据委托人提供的论文材料，委托人国家生猪种业工程技术研究中心 郑恩琴 1 篇论文收录情况如下表。

序号	论文名称	发表刊物及发表的年月卷期/页码等	作者排名	论文等级	作者文中单位	收录情况	影响因子	中科院大类分区
1	基于微课的猪生产学实验 课程混合教学研究	科教导刊（电子版） 出版年：2019 卷期：34 页码：119-120 文献号： 文献类型：教改论文	第一作者	无	国家生猪种业 工程技术研究 中心	超星期刊	无	无

说明：论文等级和中科院大类分区按《华南农业大学学位论文评阅方案（试行）》划分。

报告免责声明：如未盖章，报告无效



子课题任务合同书

(2022 年度)

子课题名称:	体型性状相关功能基因 与分子标记筛选、鉴别及应用
所属课题编号:	NK2022110301
课题牵头单位:	江西农业大学
子课题承担单位:	华南农业大学
子课题负责人:	郑恩琴
执行期限:	2022 年 6 月 30 日 至 2022 年 12 月 31 日

2022 年 6 月

基本信息表

所属课题		精准表型数据库构建及生长繁殖相关功能基因与分子标记筛选、鉴别					
所属课题编号		NK2022110301					
子课题名称		体型性状相关功能基因与分子标记筛选、鉴别及应用					
子课题承担单位	单位名称	华南农业大学					
	法定代表人	刘雅红		单位性质		高等院校	
	单位主管部门	广东省教育厅		组织机构代码		124400004554165634	
	通信地址	广州市天河区五山路 483 号		邮政编码		510642	
	单位开户名称	华南农业大学					
	银行账号	3602002609000310520		开户地点		广东省广州市	
	开户银行（全称）	中国工商银行广东省广州五山支行		银行机构代码		102581000546	
	子课题管理部门负责人	倪慧群		联系方式		02085283435	
	财务管理部门负责人	刘烨		联系方式		02085288032	
子课题负责人	姓 名	郑恩琴	性 别	女	出生日期	19810801	
	国 籍	中国	身份证号码	441322198108016041			
	工作单位	华南农业大学		职 务	高级实验师		
	最高学位	<input type="checkbox"/> 博士 <input checked="" type="checkbox"/> 硕士 <input type="checkbox"/> 学士 <input type="checkbox"/> 其他					
	职 称	<input type="checkbox"/> 正高级 <input checked="" type="checkbox"/> 副高级 <input type="checkbox"/> 中级 <input type="checkbox"/> 其他					
	电子邮箱	Eqzheng@scau.edu.cn		联系方式		13826029036	
经费估算		总经费 21.86 万元。其中，中央财政经费 21.86 万元，地方配套资金 0.00 万元，自筹资金 0.00 万元。					
子课题周期节点		起始时间	2022 年 6 月 30 日		结束时间	2022 年 12 月 31 日	
		实施周期	6 个月				
课题参加人数	共 6 人	高级职称 1 人，中级职称 0 人，初级职称 0 人，其他 5 人					
		博士学位 1 人，硕士学位 2 人，学士学位 3 人，其他 人					

填表说明：

- 1.单位名称必须与单位公章名称一致。
- 2.单位开户名称应与单位名称一致，如有开户名称不一致等特殊情况，必须提供证明文件。
- 3.子课题承担单位为企业的，银行账号须为新开设的专用一般账户。

八、子课题参加人员基本情况表

填表说明:

1. 专业技术职称: A. 正高级 B. 副高级 C. 中级 D. 初级 E. 其他;
2. 投入本课题的全时工作时间(人月)是指在课题实施期间该人总共为课题工作的满月度工作量;累计是指课题组所有人员投入人月之和;
3. 人员分类代码: A. 课题负责人 B. 课题骨干(指子课题负责人) C. 其他人员(仅填固定工作人员);
4. 工作单位: 填写单位全称, 其中高校要具体填写到所在院系。

序号	姓名	性别	出生日期	国籍	证件类型	证件号码	专业技术职称	职务	最高学位	专业	投入本课题的全时工作时间(人月)	人员分类代码	工作单位	签名
1	郑恩琴	女	19810801	中国	身份证	441322198108016041	B	高级实验师	硕士	动物遗传育种与繁殖	6	B	华南农业大学	郑恩琴
2	曹露	女	19910524	中国	身份证	500383199105240403	无	博士后	博士	动物遗传育种	6	C	华南农业大学	曹露
3	庄站伟	男	19931023	中国	身份证	41148119931023811X	无	博士研究生	硕士	动物遗传育种与繁殖	6	C	华南农业大学	庄站伟
4	吴杰	女	19960919	中国	身份证	422826199609192523	无	博士研究生	本科	动物遗传育种与繁殖	6	C	华南农业大学	吴杰
5	范振飞	男	20000101	中国	身份证	341222200001018232	无	硕士研究生	本科	动物遗传育种与繁殖	6	C	华南农业大学	范振飞
6	孟祥伦	男	19970921	中国	身份证	150429199709215091	无	硕士研究生	本科	动物遗传育种与繁殖	6	C	华南农业大学	孟祥伦
累计											36	/	/	/

九、经费预算

表 1 子课题预算支出科目表

金额单位：万元

序号	预算科目名称	金额
	(1)	(2)
1	一、中央财政资金	21.86
2	(一) 直接费用	18
3	1.设备费	0
4	其中：购置设备费	0
5	2.业务费	15
6	3.劳务费	3
7	(二) 间接费用	3.86
8	二、其他来源资金	0
9	三、合计	21.86

课题任务合同书签署

课题牵头单位（甲方）：江西农业大学

法定代表人签字：



课题负责人签字：



子课题承担单位（乙方）：华南农业大学

法定代表人签字：

子课题负责人签字：



子课题任务合同书

(2023 年度)

子课题名称:	体型性状相关功能基因 与分子标记筛选、鉴别及应用
所属课题编号:	NK2022110301
课题牵头单位:	江西农业大学
子课题承担单位:	华南农业大学
子课题负责人:	郑恩琴
执行期限:	2023 年 1 月 1 日 至 2023 年 12 月 31 日

2023 年 1 月




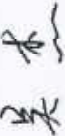
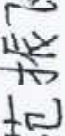

基本信息表

所属课题		精准表型数据库构建及生长繁殖相关功能基因与分子标记筛选、鉴别					
所属课题编号		NK2022110301					
子课题名称		体型性状相关功能基因与分子标记筛选、鉴别及应用					
子课题承担单位	单位名称	华南农业大学					
	法定代表人	刘雅红		单位性质		高等院校	
	单位主管部门	广东省教育厅		组织机构代码		124400004554165634	
	通信地址	广州市天河区五山路 483 号		邮政编码		510642	
	单位开户名称	华南农业大学					
	银行账号	3602002609000310520		开户地点		广东省广州市	
	开户银行（全称）	中国工商银行广东省广州五山支行		银行机构代码		102581000546	
	子课题管理部门负责人	倪慧群		联系方式		02085283435	
	财务管理部门负责人	刘烨		联系方式		02085288032	
子课题负责人	姓 名	郑恩琴	性 别	女	出生日期	19810801	
	国 籍	中国	身份证号码	441322198108016041			
	工作单位	华南农业大学		职 务	高级实验师		
	最高学位	<input type="checkbox"/> 博士 <input checked="" type="checkbox"/> 硕士 <input type="checkbox"/> 学士 <input type="checkbox"/> 其他					
	职 称	<input type="checkbox"/> 正高级 <input checked="" type="checkbox"/> 副高级 <input type="checkbox"/> 中级 <input type="checkbox"/> 其他					
	电子邮箱	Eqzheng@scau.edu.cn		联系方式		13826029036	
经费估算		总经费 32.79 万元。其中，中央财政经费 32.79 万元， 地方配套资金 0.00 万元，自筹资金 0.00 万元。					
子课题周期节点		起始时间	2023 年 1 月 1 日		结束时间	2023 年 12 月 31 日	
		实施周期	12 个月				
课题参加人数	共 6 人	高级职称 1 人，中级职称 0 人，初级职称 0 人，其他 5 人					
		博士学位 1 人，硕士学位 2 人，学士学位 3 人，其他 人					

填表说明：

- 1.单位名称必须与单位公章名称一致。
- 2.单位开户名称应与单位名称一致，如有开户名称不一致等特殊情况，必须提供证明文件。
- 3.子课题承担单位为企业的，银行账号须为新开设的专用一般账户。

八、子课题参加人员基本情况表

填表说明：1. 专业技术职称：A. 正高级 B. 副高级 C. 中级 D. 初级 E. 其他； 2. 投入本课题的全时工作时间（人月）是指在课题实施期间该人总共为课题工作的满月度工作量；累计是指课题组所有人员投入人月之和； 3. 人员分类代码：A. 课题负责人 B. 课题骨干（指子课题负责人） C. 其他人员（仅填固定工作人员）； 4. 工作单位：填写单位全称，其中高校要具体填写到所在院系。														
序号	姓名	性别	出生日期	国籍	证件类型	证件号码	专业技术职称	职务	最高学位	专业	投入本课题的全时工作时间（人月）	人员分类代码	工作单位	签名
1	郑恩琴	女	19810801	中国	身份证	441322198108016041	B	高级实验师	硕士	动物遗传育种与繁殖	10	B	华南农业大学	
2	曹露	女	19910524	中国	身份证	500383199105240403	无	博士后	博士	动物遗传育种与繁殖	6	C	华南农业大学	
3	庄站伟	男	19931023	中国	身份证	41148119931023811X	无	博士研究生	硕士	动物遗传育种与繁殖	6	C	华南农业大学	
4	吴杰	女	19960919	中国	身份证	422826199609192523	无	博士研究生	本科	动物遗传育种与繁殖	10	C	华南农业大学	
5	范振飞	男	20000101	中国	身份证	341222200001018232	无	硕士研究生	本科	动物遗传育种与繁殖	10	C	华南农业大学	
6	孟祥伦	男	19970921	中国	身份证	150429199709215091	无	硕士研究生	本科	动物遗传育种与繁殖	10	C	华南农业大学	
累计											52	/	/	/

九、经费预算

表 1 子课题预算支出科目表

金额单位：万元

序号	预算科目名称	金额
	(1)	(2)
1	一、中央财政资金	32.79
2	(一) 直接费用	27.43
3	1.设备费	0
4	其中：购置设备费	0
5	2.业务费	23.47
6	3.劳务费	3.96
7	(二) 间接费用	5.36
8	二、其他来源资金	0
9	三、合计	32.79

课题任务合同书签署

课题牵头单位（甲方）：江西农业大学

法定代表人签字：



课题负责人签字：



2023 年 10 月 8 日

子课题承担单位（乙方）：华南农业大学

法定代表人签字：

子课题负责人签字：



2023 年 10 月 8 日

子课题任务合同书

(2024—2026)

子课题名称:	体型性状相关功能基因 与分子标记筛选、鉴别及应用
所属课题编号:	NK2022110301
课题牵头单位:	江西农业大学
子课题承担单位:	华南农业大学
子课题负责人:	郑恩琴
执行期限:	2024年1月1日至2026年12月31日

2024年8月

基本信息表

所属课题		精准表型数据库构建及生长繁殖相关功能基因与分子标记筛选、鉴别					
所属课题编号		NK2022110301					
子课题名称		体型性状相关功能基因与分子标记筛选、鉴别及应用					
子课题承担单位	单位名称	华南农业大学					
	法定代表人	薛红卫		单位性质		高等院校	
	单位主管部门	广东省教育厅		组织机构代码		124400004554165634	
	通信地址	广州市天河区五山路 483 号		邮政编码		510642	
	单位开户名称	华南农业大学					
	银行账号	3602002609000310520		开户地点		广东省广州市	
	开户银行（全称）	中国工商银行广东省广州五山支行		银行机构代码		102581000546	
	子课题管理部门负责人	苏弟华		联系方式		02085283435	
	财务管理部门负责人	肖斐		联系方式		02085288032	
子课题负责人	姓 名	郑恩琴	性 别	女	出生日期	19810801	
	国 籍	中国	身份证号码	441322198108016041			
	工作单位	华南农业大学		职 务	高级实验师		
	最高学位	<input type="checkbox"/> 博士 <input checked="" type="checkbox"/> 硕士 <input type="checkbox"/> 学士 <input type="checkbox"/> 其他					
	职 称	<input type="checkbox"/> 正高级 <input checked="" type="checkbox"/> 副高级 <input type="checkbox"/> 中级 <input type="checkbox"/> 其他					
	电子邮箱	Eqzheng@scau.edu.cn		联系方式		13826029036	
经费估算		总经费 27.55 万元。其中，中央财政经费 27.55 万元， 地方配套资金 0.00 万元，自筹资金 0.00 万元。					
子课题周期节点		起始时间	2024 年 1 月 1 日		结束时间	2024 年 12 月 31 日	
		实施周期	12 个月				
课题参加人数	共 6 人	高级职称 2 人，中级职称 0 人，初级职称 0 人，其他 4 人					
		博士学位 0 人，硕士学位 3 人，学士学位 3 人，其他 0 人					

填表说明：

1.单位名称必须与单位公章名称一致。

2.单位开户名称应与单位名称一致，如有开户名称不一致等特殊情况，必须提供证明文件。

3.子课题承担单位为企业的，银行账号须为新开设的专用一般账户。

八、子课题参加人员基本情况表

填表说明： 1. 专业技术职称：A. 正高级 B. 副高级 C. 中级 D. 初级 E. 其他；
 2. 投入本课题的全时工作时间（人月）是指在课题实施期间该人总共为课题工作的满月度工作量；累计是指课题组所有人员投入人月之和；
 3. 人员分类代码：A. 课题负责人 B. 课题骨干（指子课题负责人）C. 其他人员（仅填固定工作人员）；
 4. 工作单位：填写单位全称，其中高校要具体填写到所在院系。

序号	姓名	性别	出生日期	国籍	证件类型	证件号码	专业技术职称	职务	最高学位	专业	投入本课题的全时工作时间（人月）	人员分类代码	工作单位	签名
1	郑恩琴	女	1981.08.01	中国	身份证	441322198108016041	B	高级实验师	硕士	动物遗传育种与繁殖	6	B	华南农业大学国家生猪种业工程技术研究中心	
2	刘满清	女	1977.09.20	中国	身份证	320102197709204620	B	副主任	硕士	动物遗传育种与繁殖	6	C	华南农业大学动物科学学院	
累计											12	/	/	/

课题任务合同书签署

课题牵头单位（甲方）：江西农业大学



法定代表人签字：



课题负责人签字：

（公章）

2024年11月1日

子课题承担单位（乙方）：华南农业大学

法定代表人签字：

子课题负责人签字：



（公章）

2024年11月1日

国家重点研发计划“畜禽新品种培育与现代牧场科技创新”专项

“南方区地方猪新品种（配套系）培育及良繁”项目

课题3“南方区高繁地方猪新品种培育与产业化”

任务2“高繁大黑猪新品种培育及繁育体系构建”

任务合作协议

甲方（课题3承担单位）：重庆市种猪场有限公司

通讯地址：重庆市荣昌区广顺街道高瓷村

课题负责人：郭宗义

乙方（任务2承担单位）：华南农业大学

通讯地址：广东省广州市天河区五山街道五山路483号

任务负责人：李紫聪

丙方（子任务承担单位）：华南农业大学

通讯地址：广东省广州市天河区五山街道五山路483号

子任务4负责人：郑恩琴

本协议各方共同申请国家重点研发计划“畜禽新品种培育与现代牧场科技创新”专项“南方区地方猪新品种（配套系）培育及良繁”项目，甲方为课题3“南方区高繁地方猪新品种培育与产业化”承担单位，乙方为课题3任务2承担单位，丙方为课题3任务2子任务4承担单位。甲、乙、丙三方经过平等协商，在真实、充分地表达各自意愿的基础上，根据《中

华人民共和国民法典》以及国家重点研发计划管理规定，达成如下协议，并由三方共同恪守。

第一条 任务分工

甲方作为课题 3 承担单位，承担整个课题所有内容的协调推进、过程管理、年度报告编写等工作。

乙方作为课题 3 任务 2 承担单位，积极配合，按照合作协议的规定按时完成研究计划以及科技部要求提交的课题年度报告、中期报告及验收（结题）报告等科技报告的编写。

丙方作为课题 3 任务 2 子任务 4 承担单位，积极配合，按照合作协议的规定按时完成研究计划以及科技部要求提交的课题年度报告、中期报告及验收（结题）报告等科技报告的编写。

第二条 乙方在本课题中研究任务、考核指标

1. 研究任务：

高繁大黑猪新品种培育及繁育体系构建。在前期利用粤东黑猪等地方猪资源创建育种核心群的基础上，基于繁殖性状育种策略的基因组育种技术与常规育种相结合，应用毛色基因分子标记辅助选择技术，持续选育产仔数高、产肉量高的高繁大黑猪新品种；应用杂种优势预测技术筛选杂交配套组合，集成新品种高效生产配套技术，建立从核心育种群到优质商品猪生产的高效繁育体系，开展产业化。

2. 考核指标：

（1）育成达到国家审定要求的高繁大黑猪新品种 1 个，基础母猪 1000 头以上。经产母猪总产仔数 14 头以上，达 115kg 日龄低于 210 天，育肥期料重比 2.9:1 以下，115kg 体重屠宰时瘦肉率 53%以上、肌肉脂肪含量 2.3%以上；

（2）申请专利 1 项，发表论文 1 篇，制定新品种标准 1 个；

（3）培养研究生 2 名以上。

第三条 丙方在本课题中研究任务、考核指标

1. 研究任务:

高繁大黑猪新品种培育及繁育体系构建。在前期利用粤东黑猪等地方猪资源创建育种核心群的基础上,基于繁殖性状育种策略的基因组育种技术与常规育种相结合,应用毛色基因分子标记辅助选择技术,持续选育产仔数高、产肉量高的高繁大黑猪新品种;应用杂种优势预测技术筛选杂交配套组合,集成新品种高效生产配套技术,建立从核心育种群到优质商品猪生产的高效繁育体系,开展产业化。

2. 考核指标:

(1) 育成达到国家审定要求的高繁大黑猪新品种 1 个,基础母猪 1000 头以上。经产母猪总产仔数 14 头以上,达 115kg 日龄低于 210 天,育肥期料重比 2.9:1 以下,115kg 体重屠宰时瘦肉率 53%以上、肌肉脂肪含量 2.3%以上;

(2) 申请专利 1 项;

(3) 培养研究生 1 名以上。

第四条 经费分配

本专项项目经费总额为 7400 万元,其中申请专项资金 5000 万元,申请地方财政资金 980 万元,企业自筹资金 1420 万元。本课题经费总额为 1230.00 万元(其中中央财政资金 660.00 万元,重庆市地方财政资金 330.00 万元,企业自筹资金 240.00 万元),其中乙方为 420.00 万元(全部为专项资金),其中乙方分配丙方 50 万元,经费应严格按照《国务院办公厅关于改革完善中央财政科研经费管理的若干意见》(国办发〔2021〕32 号)和《财政部 科技部关于印发〈国家重点研发计划资金管理办法〉的通知》(财科教〔2021〕178 号)等合理支出。课题、任务的设备费,业务费和劳务费科目总经费调整需报课题承担单位审批,并报项目牵头单位备案。

第五条 关于科研成果、知识产权和奖励的相关约定

在课题执行过程中涉及到的信息公开与分享、科研成果处理、知识产权申请与转让,奖励申报和收益分享等事宜按照以下约定执行,本协议未尽之处,应采取“一事一议”的方式签订补充协议:



1. 课题执行期间，各方承诺尽最大可能互为提供资料数据，共享研究成果，但相关资料和数据仅限于各方的研究目的，任何一方都不得将其他方未公开的材料和资料向其他方转移和泄露。

2. 课题承担单位与参加单位在课题执行日之前各自所获得的知识产权及相应权益均归各自所有，不因共同合作本课题而改变。

3. 在课题执行过程中，各方应对科技成果及时采取知识产权保护措施，并按照国家科技计划知识产权管理相关规定决定归属。独自完成的科技成果及获得的知识产权归各方独自所有，相关成果被授予的奖励归各方独自所有。各方共同完成的科技成果及其形成的知识产权归各方共有，共同享有知识产权使用权，相关成果获得的荣誉和奖励归完成各方共有。

4. 共有知识产权所有权申请及转让需要各方共同同意，并另行起草签署书面约定明确归属和收益共享方式。无论是独有还是共有的知识产权转让，课题各参与方有以同等条件优先受让的权利。

5. 各方承诺本项目产生的所有科学数据无条件，按期递交到科技部指定的平台，在专项约定的条件下对专项各承担单位，乃至今后面向所有的科技工作者和公众开放共享。

第六条 保密约定

各方均对对方提供的技术情报和资料等承担保密义务，不论本合作协议是否变更、解除或终止，本条款长期有效。

第七条 课题资产管理与归属约定

本课题中央财政经费支持的设备购置与样机试制，资料材料以及数据等资产按国家科技计划管理办法相关要求执行；自筹经费提供方与使用方应另行签订相关协议约定自筹经费投入产生的资产归属问题。

1、中央财政经费投入研制或购置的设备，开发的样机、资料材料及数据等按国家科技计划有关法规决定所有权和归属，并至少在课题参加单位间提供共享。

2、配套自筹经费研制或购置的设备，开发的样机，资料材料及数据等，在不违反国家法律和相关规定的基礎上，按自筹经费提供方与使用方

之间的协议处理。

第八条 课题执行违约责任

1. 根据任务书约定，各方负有按时完成各自负责的任务并达到相关考核指标的义务。如乙方和丙方研发进展滞后，甲方有权督促相关责任方加快进度；出现进展严重滞后并影响课题甚至整个项目考核指标完成的情况，甲方报项目牵头单位并报批专业机构后，有权缓拨、停拨、甚至追缴部分或全部课题经费。

2. 各方为完成任务书规定研究任务的支出，超出各自预算的部分由各方自行承担。

3. 违反本协议第四条关于科研成果和知识产权申请和权属等约定，违约方应向知识产权所属人支付违约金赔偿相关损失；在课题承担单位或有关部门调解无法达成谅解的情况下，项目各方均有权通过法律途径追究违约方责任，但相关纠纷不作为影响本课题研究进度的理由。

4. 本课题因难以克服的技术挑战或无法预见的客观条件变化而无法达到预期考核指标的情况，应及时通知甲方并报项目牵头单位，需要时应及时报告专业机构申请调整，责任和损失由各方协调共同决定承担方式。因责任方未及时通知课题承担单位造成的额外损失，由相关责任方自行承担。

5. 因不可抗力不能履行任务书规定义务时，可以免除违约责任，但应及时通知甲方并报项目牵头单位，需要时应根据相关流程及时报告专业机构。在出现不可抗力的情况下，各方均采取适当措施减轻损失。因乙方未及时采取应对措施或通知甲方造成的额外损失，由相关责任方自行承担。

第九条 课题过程管理及验收

课题承担单位应积极配合专业机构和项目牵头单位对课题执行的过程管理和验收，并采取召开会议、进展报告等方式协调和促进项目和课题的执行，督促和保证本课题的研究任务按时完成，并达到相应考核指标。

课题参加单位应严格按照国家科技计划管理的相关规定和办法执行课题预算，保证任务书规定的研究任务按时完成，并达到相应考核指标。

因一方或几方原因导致整个项目或课题验收不通过，相关参加单位应

承担责任。

第十条 补充协议签署和争议解决办法

1、若课题执行过程中，任何重大调整（如任务考核指标调整、经费调整，参加单位变化等）都应及时通知课题承担单位和项目牵头单位，报专业机构批准后签署补充协议。补充协议应对调整后的各方责任义务进行约定，与本协议具有同等效力。

2、在项目执行过程中，各参加单位发生争议应当友好协商解决。课题承担单位和项目牵头单位出面协调无法达成一致的，可请求专业机构或相关科技主管部门进行调解。协商和调解不成的，双方同意提交甲方所在地的仲裁委员会仲裁。

本协议自各方盖章（签字）之日起生效（课题承担单位加盖骑缝章），有效期至项目验收合格之日。协议一式 3 份，甲方保留 1 份，乙方保留 1 份，丙方保留 1 份。具有同等法律效力。

（此后无正文）

甲方（盖章）：重庆市种猪场有限公司

法定代表人（签章）：

课题负责人（签字）：

2022 年 1 月 20 日

乙方（盖章）：华南农业大学

法定代表人（签章）：

任务负责人（签字）：

2022年1月18日

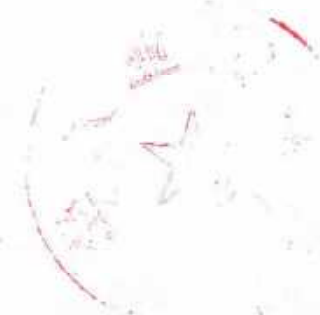
丙方（盖章）：华南农业大学

法定代表人（签章）：

子任务负责人（签字）：郑恩

2022年1月18日

2022年1月18日



**广东省重点领域研发计划项目
课题任务书**

项目名称： 种猪杂交育种技术创新与高繁优质新品系培育

项目编号： 2022B0202090002

项目负责人： 吴珍芳

课题 1 负责人： 王翀

课题 2 负责人： 张豪

课题 3 负责人： 郑恩琴

课题 4 负责人： 吴珍芳

所在单位（盖章）： 华南农业大学

项目起止年限： 2022 年 7 月 1 日至 2025 年 12 月 31 日

二〇二二年七月

一、任务分工、考核指标与经费分配

1. 课题1“特色优异种质性状和优异基因发掘”，负责人：王翀

主要研究内容：以粤东黑猪、里岔黑猪、杜洛克猪 S22 系、大白猪 W64 系等 4 个品种为研究对象，开展高繁和优质型猪种质资源鉴定评价，深度发掘繁殖能力、猪肉品质、生长速度、饲料转化率等特色优异种质性状及其优异基因，鉴定与重要经济性状相关的有效标记位点，解析高繁、优质、快长、节粮等特色优异种质性状形成的分子遗传机制和规律，建立高繁、优质、黑毛等性状的精准分子育种技术。

考核指标：系统开展猪性能评价 1 个品种（系）以上，获得特色、优质等性状紧密相关基因或选择标记 3 个以上。发表高水平论文 1 篇，申请国家专利 2 件，培养人才 2 人。

经费分配：50 万元。

2. 课题2“猪种质资源表型与基因型精准分析”，负责人：张豪

主要研究内容：以繁殖力高、肉质优良、生长速度快、饲料转化率高为核心性状，以粤东黑猪、里岔黑猪、杜洛克猪 S22 系、大白猪 W64 系为种质资源，开展大规模的表型采集、基因分型和表型-基因型精准分析，构建高繁和优质型猪种质资源“表型-基因型”数据库。

考核指标：建立重要经济性状表型-基因型数据库 1 个。发表高水平论文 1 篇，获得计算机软件著作权 1 项以上，培养人才 2 人。

经费分配：50 万元。

3. 课题3“杂交育种与现代分子设计育种技术研发”，负责人：郑恩琴

主要研究内容：构建高繁和优质型全基因组选择参考群体，建立种猪群体分子血统鉴定方法、杂种优势分子预测模型。利用基因挖掘和基因型精准分析研究成果，完善繁殖力、肉质、生长速度、饲料转化率等性状的全基因组选种、选配方法，为新品系的选育提供杂交育种与现代分子设计育种技术体系。

考核指标：建立杂交育种与现代分子设计育种技术体系 1 套。发表高水平论文 2 篇，申请国家专利 2 件，培养人才 2 人。

经费分配：100 万元。

4. 课题4“猪优良品系核心育种群组建及选育”，负责人：吴珍芳

主要内容：以优质、高繁殖力、毛色全黑，大体型为育种方向，以肌肉脂肪含量、产仔数、产活仔数等为主选性状，兼顾饲料转化率、日增重、瘦肉率，在前期采用粤东黑猪、里岔黑猪与高繁殖力大白猪 W64 进行正反杂交已获得的 F1 代杂种的基础上，采取闭锁核心群群体继代选育法，采用杂交育种结合现代分子设计育种技术，培育以优质、和高繁殖力为突出特点的“华系大黑猪”新品系。以高繁殖力、优质为育种方向，以产仔数、产活仔数、肌肉脂肪含量、饲料转化率、日增重、瘦肉率等为主选性状，在前期已开展的育种工作基础上优化组建优质型杜洛克猪 S22 系、高繁殖力大白猪 W64 系育种核心群，采取开放式核心群群体继代选育法，采用杂交育种结合现代分子设计育种技术，开展“华系”优质型杜洛克猪新品系选育、高繁殖力大白猪新品系选育。

考核指标：培育重要经济性状表型突出，达四个世代以上的猪专门化品系 1 个以上，新品系核心群体规模 500 头以上。发表高水平论文 2 篇，申请国家专利 2 件，培养人才 2 人。

经费分配：100 万元。

二、经费管理

(1) 课题负责人对经济事项支出的真实性、合法性、合理性、相关性、安全性和使用效益负责，实行“谁负责，谁把关”的原则，保证经费支出合理合规。

(2) 按照国家、省及华南农业大学科研经费管理办法对经费进行管理，每一笔报账支出登记并保留相应佐证材料，如低值易耗品验收结算单、送货单、合同、测序报告等，以备项目验收时备查。

(3) 项目结题验收时，各课题负责人需积极配合，按照审计要求提供所需材料。

三、其它

1. 课题负责人按照约定的任务和考核指标开展工作，在项目结题验收前，按照任务分工完成研究任务和考核指标。

2. 在项目实施过程中，项目主持人有权检查任务进度。

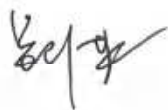
3. 课题负责人须配合提供和项目任务相关的佐证材料。

4. 未尽事宜各方协商解决。

5. 本协议一式四份，自双方签字之日起生效。

四、任务书签署

项目负责人签字：



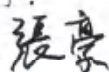
日期：2022.7.4

课题1负责人签字：



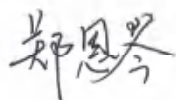
日期：2022.7.4

课题2负责人签字：



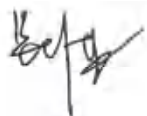
日期：2022.7.4

课题3负责人签字：



日期：2022.7.4

课题4负责人签字：



日期：2022.7.4

项目所在单位盖章（学校公章）



受理编号: c24140500000413

项目编号: 2024A1515010967

文件编号: 粤基金字(2024)7号

广东省基础与应用基础研究基金项目

任务书

项目名称: 猪12号染色体影响初生重主效QTL的分子遗传机理解析

项目类别: 广东省自然科学基金-面上项目

项目起止时间: 2024-01-01 至 2026-12-31

管理单位(甲方): 广东省基础与应用基础研究基金委员会

依托单位(乙方): 华南农业大学

通讯地址: 广东省广州市天河区五山路483号

邮政编码: 510642

单位电话: 020-85283435

项目负责人: 郑恩琴

联系电话: 02085280369



(广东科技微信公众号)



(查看任务书信息)



(受理纸质材料二维码)

广东省基础与应用基础研究
基金委员会
二〇二〇年制

一、主要研究内容和要达到的目标

主要内容：初生重对仔猪断奶前存活率及育肥猪生长速度具有重要影响，直接关系到能繁母猪的生产效率。初生重过低，会导致仔猪断奶前死亡率升高，同时也会对猪育肥期生长速度造成负面影响，不仅给养猪生产造成较大的直接经济损失，而且间接影响规模化猪场的“全进全出”批次化生产。因此，在保障能繁母猪产仔数的情况下，减少低初生重仔猪的出现频率，是提升我国能繁母猪生产效率的关键，也是保障我国生猪产业产能升级的重大需求。本项目拟采用整合基因组学、转录组学、蛋白质组学及染色质开放区定位的多组学研究方法解析该QTL，预期通过扩充不同品系的大白群体，开展共享单倍型分析，精细定位该主效QTL的置信区间；进而利用多组学策略鉴别相应的主效基因及大效应因果突变，所鉴别的突变将可以直接应用于猪初生重的早期分子标记辅助选择。该研究不仅能为仔猪产前遗传发育提供新的视角，而且可产生能够实际应用的分子育种标记，研究内容符合产业的迫切需要，具备重大的产业应用价值。

要达到的目标：

- 1) 通过增加不同品系大白猪的遗传素材，将SSC12上影响仔猪初生重的QTL的置信区间缩小至原定位区间的一半左右（约200Kb）；
- 2) 通过整合基因组、转录组、蛋白质组及候选突变的功能注释信息，鉴别SSC12上影响仔猪初生重的因果基因和因果突变位点；
- 3) 通过后续细胞功能实验，对因果基因及其因果突变位点进行功能验证，揭示其影响仔猪初生重的分子遗传机理，进而为仔猪初生重的早期分子选育提供参考。

二、项目预期获得的研究成果及形式

论文及专著情况	国家统计源刊物以上刊物 发表论文（篇）		2		科技报告（篇）		1	
	其中被SCI/EI/ISTP收录 论文数（篇）		2		培养人才（人）		3	
	专著（册）		0		引进人才（人）		0	
专利情况(项)	发明专利		实用新型专利		外观设计专利		国外专利	
	申请	授权	申请	授权	申请	授权	申请	授权
	2	1	0	0	0	0	0	0

三、项目进度和阶段目标

(一) 项目起止时间： 2024-01-01 至 2026-12-31		
(二) 项目实施进度及阶段主要目标：		
开始日期	结束日期	主要工作内容
2024-01-01	2024-12-31	完成不同品系（丹系或英系）大白猪初生重样本的收集，在QTL初步区域内均匀选择30个标签SNP，采用MassArray飞行质谱完成实验样本的基因分型，通过关联分析和基于后裔同源的共享单倍型分析精细定位QTL。完成不同阶段实验样本的四种组织的RNA提取，并送测序公司进行转录组测序或全转录组测序，完成RNA-seq及全转录组测序数据的生物信息学处理；完成精细定位区间内关键功能基因的筛查，开展eQTL定位、QTT分析；综合eQTL定位、QTT分析及基因调控网络构建，鉴别影响仔猪初生重的因果基因。
2025-01-01	2025-12-31	开展ATAC-seq试验，获得不同组织全基因组范围的调控序列注释数据集；利用ATAC-seq数据和非编码RNA数据，结合突变在猪、人和小鼠等物种中的功能注释结果，鉴定突变是否位于因果基因的启动子区域或基因编码区。
2026-01-01	2026-12-31	采用双荧光素酶报告实验进行启动子活性分析；通过RNA干扰、基因过表达及Western blot等多种实验手段验证候选突变是否直接影响基因的转录和翻译水平，进而鉴别出影响仔猪初生重的因果突变位点。

四、项目总经费及省基金委经费预算

1. 省基金委经费下达总额： （大写）壹拾伍万圆整；（小写 ）15万元；					
2. 省基金委经费年度下达计划：					
年度	2024 年	年	年	年	年
经费(万元)	15.00				

2024A1515010967

五、人员信息

项目负责人								
姓名	证件号码	年龄	性别	职称	学历	在项目中承担的任务	所在单位	签名
郑恩琴	441322198108016041	43	女	高级实验师	硕士研究生	项目负责人	华南农业大学	郑恩琴

项目组主要成员								
姓名	证件号码	年龄	性别	职称	学历	在项目中承担的任务	所在单位	签名
杨杰	430721198706211333	37	男	副研究员	博士研究生	多组学分析	华南农业大学	杨杰
吴杰	422826199609192523	28	女	未取得	本科	转录组分析	华南农业大学	吴杰
李雪华	370725199812070689	26	女	未取得	本科	分子实验	华南农业大学	李雪华
杨莹山	370481200002263833	24	男	未取得	硕士研究生	转录组分析	华南农业大学	杨莹山
孟祥伦	150429199709215091	27	男	未取得	硕士研究生	基因分型和基因组分析	华南农业大学	孟祥伦
范振飞	341222200001018232	24	男	未取得	硕士研究生	基因分型和基因组分析	华南农业大学	范振飞

六、工作分工及财政经费分配

承担/参与单位名称 (盖章)	工作分工	省级财政科技资金分配 (万元)
华南农业大学	采用共享单倍型分析精细定位该QTL；进一步利用猪胚胎发育不同时期的转录组数据，开展表达QTL定位及基因表达量与表型关联的QTT分析，并构建与猪胚胎发育相关的基因调控网络；结合非编码RNA及染色质开放区定位等结果，初步鉴别影响猪初生重的主效基因及关键突变，并通过细胞水平实验对关键突变进行验证	15.00
	合计	15.00

八、本任务书签约各方

管理单位（甲方）：

广东省基础与应用基础研究基金委员会（盖章）



法定代表人（或法人代理）：

曾路

（签章）

2024 年 05 月 22 日

依托单位（乙方）：华南农业大学

法定代表人（或法人代理）：薛红卫

薛红卫

（盖章）

联系人（项目主管）姓名：倪慧群

Email: kjcgxk@scau.edu.cn

电话：020-85283435 / 15920301530

开户单位名称：华南农业大学

开户银行名称：广东广州工行五山支行

开户银行账号：3602002609000310520



2024 年 5 月 23 日

联系人（项目负责人）姓名：郑恩琴

郑恩琴

（签名）

Email: eqzheng@scau.edu.cn

电话：02085280369

2024 年 5 月 23 日

畜禽育种国家重点实验室 开放课题合同书

课 题 名 称：游离氨基酸对猪肉风味影响及其在肉
品内的形成机制

项 目 编 号：2021GZ04

起 止 年 限：2021.05—2022.06

课 题 负 责 人：郑恩琴

负责人所在单位：华南农业大学

通讯地址及邮编：广州天河区五山路 483 号 48 栋 618

联 系 电 话：13826029036

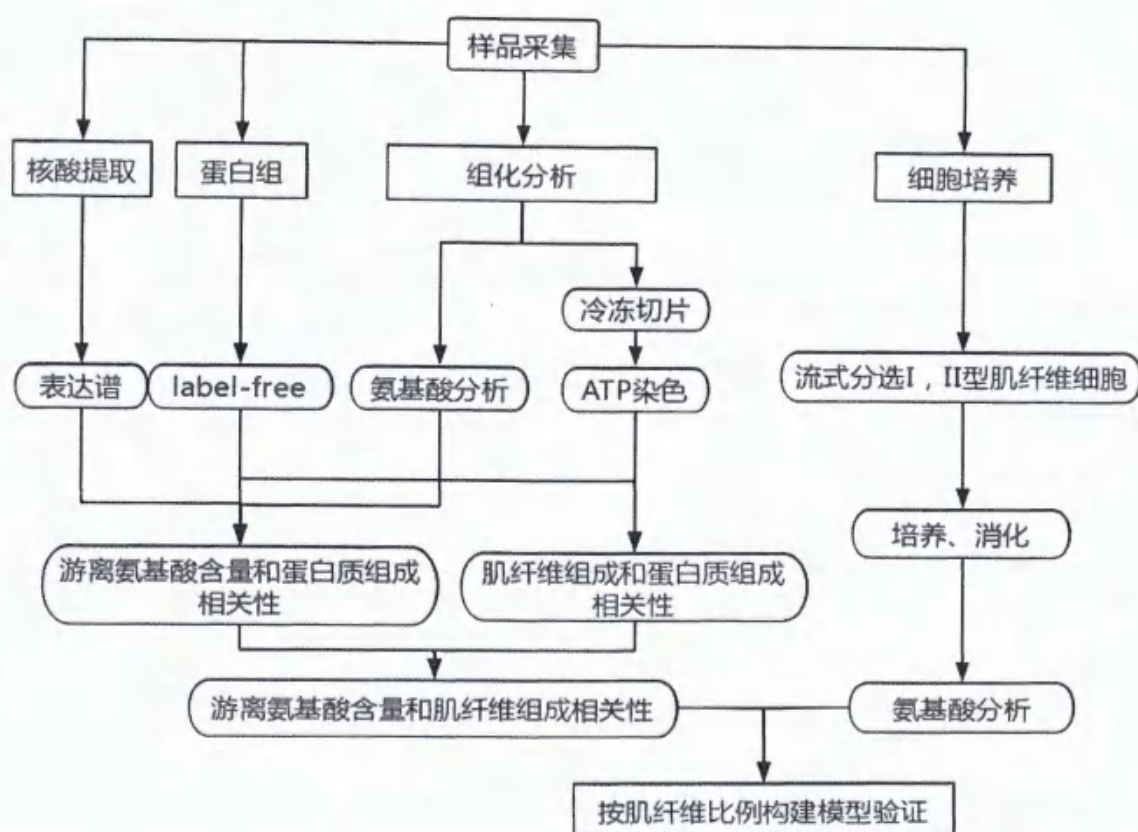
签 订 时 间：2021.6

畜禽育种国家重点实验室
广东省农业科学院动物科学研究所（畜牧研究所）

一、项目内容与目标					
(一) 主要研究内容					
<p>本项目以代谢组、蛋白组为研究手段,通过测定地方猪和瘦肉型猪肉品内游离氨基酸含量,拟建立游离氨基酸与肉品风味的量化关系,研究肉品内游离氨基酸形成的生理生化机制,为进一步研究肉质性状的遗传调控提供依据。</p> <p>主要研究内容:</p> <ol style="list-style-type: none"> 1. 测定瘦肉型猪和地方猪种生肉和熟肉游离氨基酸含量。 2. 瘦肉型猪和地方猪种肌肉蛋白组成分析。 3. 瘦肉型猪和地方猪种骨骼肌纤维类型分析。 4. 猪肌纤维类型与游离氨基酸含量相关性研究。 					
(二) 计划达到的目标					
<ol style="list-style-type: none"> 1. 建立游离氨基酸种类、含量与肉品风味的量化关系模型1个。 2. 确定肉品内与游离氨基酸含量密切相关的表型性状1个。 3. 申请专利1项,发表论文1篇。 					
(三) 将提供的研究开发成果及形式					
成果形式		成果数量	成果形式		成果数量
专利	申请	1	引进人才(人)		0
	授权	0	培养人才(人)		0
新产品(或信材料、新装备、新品种(系))		0	论文论著(篇)		1
新工艺(或新方法、新模式、新技术)		0	其中:被收录论文数	SCI	0
技术标准制定	牵头(个)	0		EI	0
	参与(个)	0			

二、项目实施方案

(一) 技术路线或实施方案



(二) 工作计划与进度安排

开始时间	结束时间	主要工作内容
2021.05	2021.07	收集相关研究信息，设计实验流程，采集样品，确定基本研究纲要。
2021.08	2022.03	按照计划安排进行实验内容，整理修改阶段性研究结果。
2022.04	2022.06	课题成果整理，论文的撰写与修改。

三、人员信息					
姓名	证件号码	学历	职称	工作分工	所在单位
郑恩琴	441322198108 016041	硕士	高级实 验师	项目实 施、统筹	华南农业大学动物 科学学院
李颖	152630199706 237015	学士	在读研 究生	蛋白组、 组化分析	华南农业大学动物 科学学院
霍梦飞	410426199611 153093	学士	在读研 究生	细胞培养 验证模型	华南农业大学动物 科学学院
孟繁明	110106198011 133913	博士	助理研 究员	相关性分 析	广东省农业科学院 动物科学研究所
辛海云	370921198805 034242	博士	助理研 究员	氨基酸分 析	广东省农业科学院 动物科学研究所
四、经费预算					
甲方核定用于乙方开放课题研究的经费 <u>8</u> 万元，供乙方进行课题研究时使用。经费使用必须专款专用、专账核算。					
经费可用于材料费、测试化验加工外协费、燃料动力费、差旅费、会议费、国际合作与交流费、出版/文献/信息传播/知识产权事物、劳务费、人员费、专家咨询费、小型设备费等。不能用于基建费。其中设备费比例不能超过 15%，其余科目不设比例限制。					
五、合同条款					
1. 甲方与乙方根据《中华人民共和国合同法》及国家有关法规和规定，为顺利完成 <u>游离氨基酸对猪肉风味影响及其在肉品内的形成机制</u> 项目(2021GZ04)，经协商一致，特订立本合同，作为甲乙双方在项目实施管理过程中共同遵守的依据。					
2. 乙方应按计划与本重点实验室合作开展课题研究，甲方为该课题研究提供必要的支持。课题到期后乙方须向甲方提交结题报告，并办理相关的项目结题手续。					

3. 在合同执行过程中,任何一方不得擅自修改合同内容,如确需修改,应由甲方乙方共同商定,重新签定合同方能生效。

4. 甲方乙方对相关技术资料负有保密责任。

5. 乙方在开放课题研究中形成的论文、专著、成果评议鉴定资料以及其他知识产权,必须标注“畜禽育种国家重点实验室”为第一完成单位。成果所有权归甲乙双方。

6. 本合同一式四份,双方各执两份,具有同等效力。签字、盖章后即生效,有效期至项目结题后一年内。各方均应负合同的法律责任,不应受机构、人事变动的影响。

7. 本合同未尽事宜双方协商解决。

六、签约各方

下达单位(甲方): 畜禽育种国家重点实验室(盖章)

实验室主任(签字):

联系电话: 020-61368882

2021年7月9日

承担单位(乙方):

承担单位联系人(签字): 李绍云

联系电话: 85284833

开户单位名称: 华南农业大学

开户银行名称: 中国工商银行广州五山支行

开户银行帐号: 3602002609000310520

2021年6月28日

项目负责人(签字): 郑恩琴

联系电话: 13826029036

2021年6月28日

项目下达文号: 江科〔2024〕215 号
项目编号: _____

江门市科技计划项目 合 同 书

项目名称: 2024年度特色农产品种养技术提升与推广

专项资金类别: 省科技专项资金 (“大专项+任务清单”)

业务类型: 江门市2024年科技支撑“百千万工程”项目

项目起止时间: 2024年09月01日 至 2025年08月31日

管理单位(甲方): 江门市科学技术局

承担单位(乙方): 广东省农业科学院蔬菜研究所

项目负责人: 郭巨先

联系电话: 13660674941

项目联系人: 胡敏春

联系电话: 020-38469427

江门市科学技术局
二〇二四年制

一、项目基本情况表

项目名称	2024年度特色农产品种养技术提升与推广			
项目起止时间	2024年09月01日 至 2025年08月31日			
项目总经费预算	4.0000万元	申请经费资助	4.0000万元	
项目摘要 (200字以内)	以特色农产品种养为核心，在现有的品种以及生产模式的基础上，通过科技帮扶引进新品种和新技术，开展科技培训让农户更快更好掌握新技术，实现种养殖产业的品种结构优化与技术的更新。在现有的农业龙头或专业合作社建立示范基地，以点带面，带动周边农户，促进农业增产农民增收。			
关键字	农产品 种植 养殖			
项目技术情况	技术领域	农业-优良动植物新品种	技术来源	自有技术
	研究活动类型	科技服务	项目研究阶段	产业化
项目已受财政资金资助情况	<div> <input type="checkbox"/> 国家（部委）财政资金资助 <input type="checkbox"/> 省级财政科技资金资助 </div> <div> <input type="checkbox"/> 省级财政其他部门资金资助 <input type="checkbox"/> 地市财政资金支持 </div> <div> <input checked="" type="checkbox"/> 其他 <input type="checkbox"/> 无 </div> <div> 简要说明（限100字）： 无 </div>			
其它需要说明的问题				

本项目是否涉及实验动物： ☐是 ☒否

二、项目研发内容和关键技术

(1) 主要研究内容：根据镇实际发展需求，可整合广东省农业科学院和华南农业大学不同学科的专家对该镇开展全方位的科技服务。以本镇的特色农产品种养为核心，在现有的品种以及生产模式的基础上，通过科技帮扶引进新品种和新技术，开展科技培训让农户更快更好掌握新技术，实现种养产业的品种结构优化与技术的更新。在现有的农业龙头或专业合作社建立示范基地，以点带面，带动周边农户，促进农业增产农民增收。

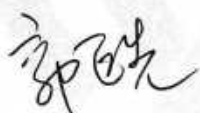
(2) 拟解决的关键问题：筛选适宜本地发展的特色新品种，建立示范基地。

(3) 创新点：本团队由广东省农业科学院和华南农业大学的专家共同组成，可以整合两个单位的优势，利用本团队最新成果帮扶当地农业产业。



三、项目考核指标

1. 本项目完成后提供的研发成果及形式				
成果形式		成果数量	成果形式	成果数量
发明专利	申请	0	新工艺(或新方法、新模式、新技术)	0
	授权	0	新产品	0
实用新型专利	申请	0	新材料	0
	授权	0	新装备	0
外观设计专利	申请	0	软件著作权	0
	授权	0	论文论著	0
国外专利	申请	0	技术标准制定	0
	授权	0	其他	0
其他成果及形式说明:				
1、引进新品种5个以上，筛选1-3个适宜当地发展的特色新品种。 2、建立1个蔬菜良种引进与示范基地，引进的新优蔬菜新品种在基地示范种植。 3、开展现场观摩会、技术培训会1-2场。				
2. 本项目完成后预计技术指标情况				
1、引进新品种5个以上，筛选1-3个适宜当地发展的特色新品种。 2、建立1个蔬菜良种引进与示范基地，引进的新优蔬菜新品种在基地示范种植。 3、开展现场观摩会、技术培训会1-2场。				

3. 本项目完成后预计社会效益情况	
新引进品种每亩农民增收500元以上	
4. 本项目完成后预计经济效益情况（单位：万元）	
累计新增销售收入	0
累计新增缴税	0
累计新增利润	0
累计新增出口创汇	0
项目负责人（签章）：  2024 年 12 月 17 日	

江门市科学技术局

四、进度和阶段目标



起止时间	主要工作内容
2024年9月 至 2025年8月	建立1个蔬菜良种引进与示范基地，引进的新优蔬菜新品种在基地示范种植。开展现场观摩会、技术培训会。
年 月 至 年 月	
年 月 至 年 月	
年 月 至 年 月	
年 月 至 年 月	
年 月 至 年 月	
年 月 至 年 月	
年 月 至 年 月	
年 月 至 年 月	

江门市科学技术局

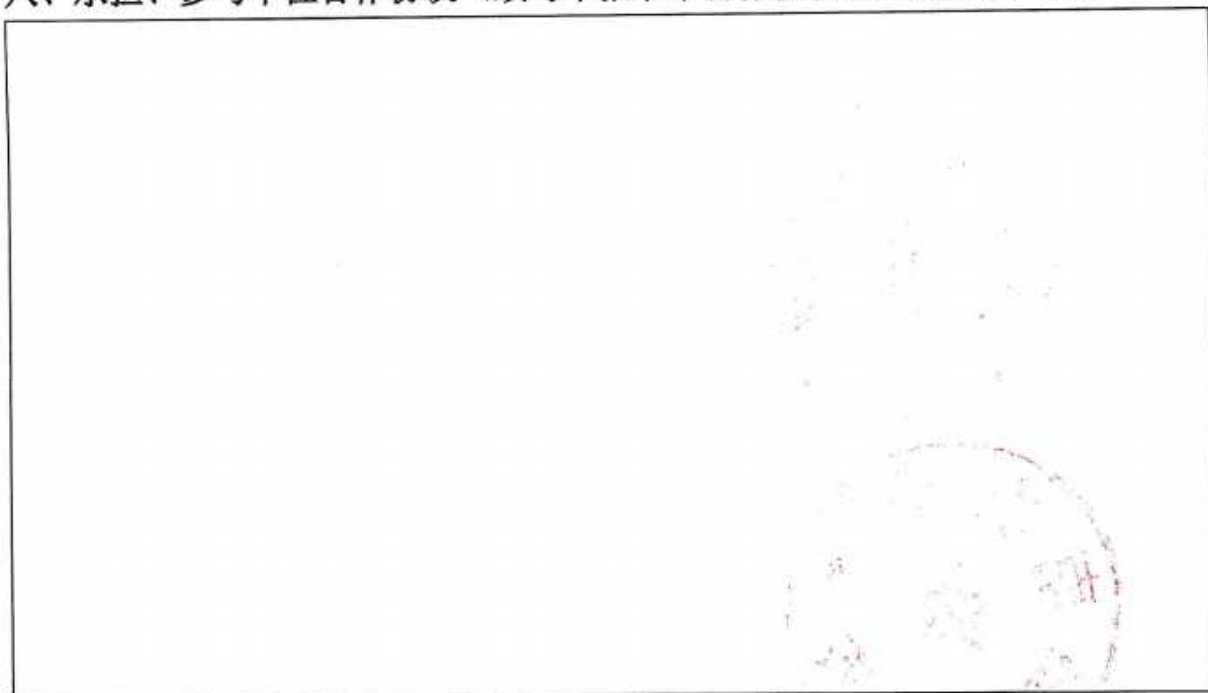
五、经费情况

经费筹集情况:						(单位: 万元)
总投入经费: 4.0000万元						
	市科技局 经费	其他资金				合计
		自有资金	贷款	地方政府 配套	其他	
已投入经费	0	0	0	0	0	0
新增经费	4	0	0	0	0	4
新增经费预算 (单位: 万元)						
		新增经费总额		其中: 市科技局经费		
支出经费	经费额	用途说明		经费额	用途说明	
基建费	0			—	—	
(一) 直接费用						
1. 设备费	0			0		
2. 材料费	0.8	购买新品种新型肥料等生产资源		0.8	购买新品种新型肥料等生产资源	
3. 测试化验加工外协费	0			0		
4. 燃料动力费	0			0		
5. 差旅费	2.5	专家去基地的差旅费		2.5	专家去基地的差旅费	
6. 会议费	0			0		
7. 学术交流费	0			0		
8. 出版/文献/信息传播/知识产权事务费	0			0		
9. 租赁费	0			0		
10. 人员费	0			0		
11. 专家咨询费	0.24	培训的专家咨询费		0.24	培训的专家咨询费	
12. 直接费其他支出	0			0		
(二) 间接费用						
管理费	0.26	管理费		0.26	管理费	
绩效	0.2	绩效		0.2	绩效	
合计	4			4		
备注						

七、承担单位及参与单位分工及经费分配情况

承担/参与单位名称 (盖章)	工作分工	新增经费 分摊(万元)	市科技局经费 分配(万元)
 广东省农业科学院蔬菜研究所	制定项目计划和实施方案,做好团队成员的分工,统筹团队各项工作,搭建团队与镇政府、驻镇工作队及科技部门的有机联系;根据镇产业发展实际,引进适宜发展的特色新品种,开展相关技术培训及现场技术指导,协助制订相关生产技术规程;对具有基础的农业龙头企业进行重点帮扶,打造一个新品种新技术引进与产业示范基地。	2.67	2.67
 华南农业大学	负责畜牧水产养殖方面的科技帮扶,包括特色新品种引进、污水处理、养殖技术指导与培训与推广等。	1.33	1.33
合计		4	4

八、承担、参与单位合作协议（须与申报书中合作协议或意向书相一致）

A large rectangular box intended for the agreement text. On the right side of the box, there is a faint, circular red stamp, likely an official seal or stamp from the relevant authority.

江门市科学技术局

九、合同条款

第一条	甲方与乙方根据《中华人民共和国民法典》及国家有关法规和规定，依据项目立项通知，为顺利完成（2024）年 2024年度特色农产品种养技术提升与推广项目（文件编号：江科〔2024〕215号），经协商一致，特签署本合同。
第二条	《项目申报书》和申报指南是本合同书填报的重要依据，合同书填报时不得降低考核指标，不得自行对主要研究任务做出调整。《项目申报书》、申报指南和本合同书将共同作为项目过程管理、验收结题和监督评估的重要依据。
第三条	甲方实行科技计划“放管服”改革，建立基于信任的管理制度，但有权按照有关规定采取日常监管、随机抽查、专项检查、中期评估、财务审计等方式对项目实施监督，并严格以本合同书中约定的任务、期限、目标和验收指标等具体内容作为监督依据。项目到期后，甲方依据本合同书对项目实施结果进行验收。甲方可根据有关规定对乙方进行科技计划信用管理。
第四条	乙方要建立以诚信为原则的自主管理制度，按如下要求执行项目： 1. 承担项目的核心研究任务。 2. 统筹协调做好资源分配和任务分工工作，履行项目组织实施和资金使用等方面的主体责任，对项目实施目标和财政资金绩效负责。 3. 完善单位内部控制制度和单位间监督制约机制。 4. 乙方负责协调落实项目自筹经费及有关保障条件，按照任务分工、任务量和时间进度合理分配和拨付财政资金，确保财政资金使用的安全有效，并签订单位间的合作协议。 5. 乙方及相关参与单位均应对财政资金实行专款专用，单独列账，并积极配合甲方（或委托专业机构等）进行的监督检查。 6. 乙方须积极配合甲方组织的评估检查；项目完成后，应主动申请验收结题。 7. 项目负责人及主要研究开发成员应实质性参与项目组织实施，不得出现挂名现象。
第五条	在本合同履行过程中，根据实际需要，乙方可按规定对项目起止时间、项目经费使用（包括自筹经费、经费分配及经费支出预算等）、项目内容、技术路线、项目参与单位、项目负责人和项目组成员等进行申请变更。
第六条	在履行本任务过程中，乙方及参与单位必须恪守科研道德准则、科研活动规范和科研诚信规定，严格遵守有关法律法规。涉及医学、生物技术和人工智能等敏感领域研究，应当按照国家有关规定设立伦理委员会，开展研究项目的伦理审查，遵循国际公认的科研伦理规范和生命伦理准则。
第七条	各方应充分理解本合同书的内容并自愿签署本合同书。未尽事宜，协商解决或签订补充协议进一步明确，甲方拥有最终解释权。
第八条	本合同一式4份，甲乙双方各留存2份。本合同自签字并加盖公章之日起生效，双方均应负本合同的法律责任，不应受机构、人事变动的影响。

十、本合同签约各方

管理单位（甲方）：江门市科学技术局

（盖章）

法定代表人（或法人代理）：刘长虹

（签章）

项目主管（联系人）：吴知豪

（签章）

电话：0750-8220818

传真：0750-8220800

E-mail: kjjzhghk@jiangmen.gov.cn

2024年12月30日

年 月 日

承担单位（乙方）：广东省农业科学院蔬菜研究所

（盖章）

法定代表人（或法人代理）

（签章）

联系人（项目主管）姓名：

李海云

（签章）

电话：18611811217

传真：02038469565

E-mail: 751817334@qq.com

地址：广东省天河区金颖路66号

乙方开户单位名称：广东省农业科学院蔬菜研究所

开户银行：中国工商银行五山支行

开户账号：3602002609000234551-0000000001

2024 年 12 月 17 日



项目批准号	31972540
申请代码	C170301
归口管理部门	
依托单位代码	51064208A0499-0932



国家自然科学基金委员会 资助项目计划书

资助类别：面上项目

亚类说明：

附注说明：

项目名称：鉴别猪13号染色体影响悬蹄过度生长的目的基因及其关键突变

直接费用：59万元 执行年限：2020.01-2023.12

负责人：杨杰

通讯地址：广东省广州市天河区五山路483号华南农业大学动物科学学院616室

邮政编码：510642 电 话：02085283780

电子邮件：jieyang2012@hotmail.com

依托单位：华南农业大学

联系人：倪慧群 电 话：020-85280070

填表日期：2019年08月17日

国家自然科学基金委员会制



简表

申请者信息	姓 名	杨杰	性 别	男	出生年月	1987年06月	民 族	汉族
	学 位	博士			职称	副研究员		
	是否在站博士后	否			电子邮件	jieyang2012@hotmail.com		
	电 话	02085283780			个人网页			
	工 作 单 位	华南农业大学						
	所 在 院 系 所	动物科学学院						
依托单位信息	名 称	华南农业大学					代码	51064208A0499
	联 系 人	倪慧群			电子邮件	kyc.jhk@scau.edu.cn		
	电 话	020-85280070			网站地址	http://kjc.scau.edu.cn/		
合作单位信息	单 位 名 称							
项目基本信息	项 目 名 称	鉴别猪13号染色体影响悬蹄过度生长的目的基因及其关键突变						
	资 助 类 别	面上项目				亚 类 说 明		
	附 注 说 明							
	申 请 代 码	C170301:猪遗传育种				C170101:畜禽性状形成的基础		
	基 地 类 别	国家生猪种业工程技术研究中心						
	执 行 年 限	2020.01-2023.12						
	直 接 费 用	59万元						



项目摘要

中文摘要:

由悬蹄过度生长导致的能繁母猪被提前淘汰,是造成我国能繁母猪生产效率低下的重要原因。然而,尽管这一现象普遍存在于瘦肉型猪种中,但国际上尚无有效的早期选育方法,迫切需要建立高效精准的分子育种技术。申请人前期通过全基因组关联分析,首次在猪13号染色体上发现一个显著影响悬蹄长度的基因座(QTL, $P < 8E-9$),其置信区间为1.2 Mb。本研究拟在此基础上,通过深度重测序、基因型填充及共享单倍型分析,精细定位该QTL;进一步利用实验群体的转录组数据,开展表达QTL定位和基因表达量与表型关联的QT T分析,并构建与目标性状相关的基因调控网络;进一步结合非编码RNA及染色质开放区定位等结果,初步鉴别影响悬蹄长度的目的基因及其关键突变;最后,通过细胞水平实验对关键突变进行验证。预期结果不仅将深入揭示趾蹄疾患的遗传本质,还将为能繁母猪提供悬蹄过度生长的早期分子育种手段,从而有效地提升我国能繁母猪的生产效率。

Abstract:

Overgrowth of dewclaws leads to the early elimination of fertile sows, which is an important reason for the low productivity of sows in China. However, although this phenomenon is prevalent in commercial pigs, there is no effective method for early culling of affected individuals. Therefore, it is urgent to establish molecular diagnostic tool to eradicate this genetic defect. Recently, using genome-wide association analysis (GWAS), we mapped a major quantitative trait loci (QTL) for the length of dewclaws on chromosome 13 in a Large White population ($P < 8E-9$). We have defined this QTL into a confidence interval of 1.2 Mb. Based on these results, this project will first refine the location of this QTL via an integrative analysis including deep resequencing, genotype imputation and identical-by-decent (IBD) mapping. Then, RNA sequencing (RNA-seq), expression QTL (eQTL), quantitative trait transcript (QTT) and the gene network analysis would be conducted. Next, the non-coding RNA and ATAC-seq data will be incorporated to identify the responsible genes and critical mutations for the length of dewclaws in pigs. Finally, the biological function of the causative mutations will be verified at cell levels. This project will not only reveal the genetic architecture of overgrowth of dewclaws in pigs, but also provide molecular diagnostic tool for early culling of individuals with overgrowth of dewclaws, thus improving sow production efficiency in China.

关键词(用分号分开): 猪; 多组学; 悬蹄过度生长; 因果基因; 因果突变

Keywords(用分号分开): pig; multi-omics; overgrowth of dewclaws; causative genes; causative mutations



项目组主要成员

编号	姓名	出生年月	性别	职称	学位	单位名称	电话	证件号码	项目分工	每年工作 时间 (月)
1	杨杰	1987.06	男	副研究员	博士	华南农业大学	02085283780	430721198706211333	项目负责人	9
2	吴珍芳	1970.03	男	教授	博士	华南农业大学	02085280369	420111197003145016	统筹项目所需样本及悬蹄表型收集	4
3	郑恩琴	1981.08	女	高级实验师	硕士	华南农业大学	02085280369	441322198108016041	不同遗传来源大白猪的组织样本搜集	6
4	黄思秀	1982.08	女	助理研究员	硕士	华南农业大学	02085280369	360424198208150042	协助DNA提取及病理切片实验	6
5	丁荣荣	1992.04	男	博士生	硕士	华南农业大学	02085283780	320682199204295430	全基因组关联分析、eQTL定位及QTL分析	8
6	庄站伟	1993.10	男	硕士生	学士	华南农业大学	02085283780	41148119931023811X	缺失基因型填充及QTL精细定位	10
7	李绍云	1989.02	女	硕士生	学士	华南农业大学	02085283780	430721198902244001	ATAC-seq实验及细胞功能实验	10
8	彭珑珑	1996.02	男	硕士生	学士	华南农业大学	02085283780	33088119960212713X	基因网络及全转录组分析	10
总人数		高级			中级	初级	博士后		博士生	硕士生
8		3			1	0	0		1	3



国家自然科学基金项目直接费用预算表（定额补助）

项目批准号：31972540

项目负责人：杨杰

金额单位：万元

序号	科目名称	金额
1	项目直接费用合计	59.0000
2	1、设备费	0.0000
3	(1)设备购置费	0.00
4	(2)设备试制费	0.00
5	(3)设备升级改造与租赁费	0.00
6	2、材料费	6.5000
7	3、测试化验加工费	34.6300
8	4、燃料动力费	0.00
9	5、差旅/会议/国际合作与交流费	5.9000
10	6、出版/文献/信息传播/知识产权事务费	4.20
11	7、劳务费	7.3600
12	8、专家咨询费	0.4100
13	9、其他支出	0.00



预算说明书（定额补助）

（请按照《国家自然科学基金项目预算表编制说明》的有关要求，对各项支出的主要用途和测算理由，以及合作研究外拨资金、单价 ≥ 10 万元的设备费等内容进行必要说明。）

项目直接费用合计：59.00 万元

1、设备费：无

2、材料费：6.5 万元

（1）分子生物学试剂盒：合计5.83万元

- a) DNA提取试剂盒（Omega，50次），720元/盒 $\times 10$ ，计0.72万元；
- b) RNA提取试剂盒（Qiagen，100次），4700元/盒 $\times 1$ ，计0.47万元；
- c) PCR产物纯化试剂盒（Omega，200次），700元/盒 $\times 3$ ，计0.21万元；
- d) 普通Taq酶（博彩）0.2万元（100元/支 $\times 20$ 支）；dNTP 0.18万元（90元/支 $\times 20$ ）；小计0.38万元；
- e) 反转录试剂盒0.35万元（1750元/盒 $\times 2$ ）；Sybr Green定量染色试剂盒0.38万元（1900元/盒 $\times 2$ ）；小计0.73万元；
- f) Western blot蛋白表达抗体，3000元/个，目的基因2个、看家基因1个，计0.9万元；Western blot检测试剂盒0.3万元（3000元/盒 $\times 1$ ）；小计1.2万元；
- g) RNA干扰实验试剂盒（RoChe）7600元/盒 $\times 1$ ；计0.76万；
- h) 功能验证实验所需试剂：胎牛血清（Gibco, 500ml）3000元/瓶 $\times 1$ ，计0.3万元；DMEM细胞培养基（Gibco, 500ml）500元/瓶 $\times 3$ ，计0.15万元；Penicillin- Streptomycin双抗（Gibco, 25ml）300元/瓶 $\times 1$ ，计0.03万元；GLUMAX缓冲液（25ml），600元/瓶 $\times 1$ ，计0.06万元；PBS（500ml），300元/瓶 $\times 3$ ，计0.09万元；荧光报告检测试剂盒（Promega, 100 assays），2000元/盒 $\times 2$ ，计0.4万元；Lipofectamine 2000转染试剂（Life, 0.75ml）0.33万元；小计1.36万元。

（2）常规试剂和实验耗材：DNA Marker、氯仿、异戊醇、无水乙醇、液氮、干冰、取样管、移液枪头、96孔板等，合计0.67万元。

3、测试化验加工费：34.63 万元

（1）基因组重测序：30 个样本基因组深度重测序，每个样本建库 150 元、测序 54G 需 3240 元，合计 $(150+3240) \times 30=10.17$ 万元；

(2) 转录组测序：常规 RNA-seq 测序 100 个样，每个样本建库 350 元、测序 6G 需 360 元，共计 $(350+360) \times 100 = 7.1$ 万元；全转录测序 10 个样本，共 3.8 万元。合计 10.9 万元；

(3) 精细定位中候选位点的 MassArray 飞行质谱实验：两板（384 孔）个体，每板可检测 30 个 SNP 位点， $1.2 \text{ 万元} \times 2 = 2.4$ 万元；

(4) 基因调控元件注释：12 个样本的 ATAC-seq 测序，每个样本建库 3000 元，测序 12G 需 4300 元，合计 $(3000+4300) \times 12 = 8.76$ 万元；

(5) 基因功能研究各种实验所需 DNA 一代测序费用，2.4 万元。

4、燃料动力费：无

5、差旅/会议/国际合作与交流费：5.9 万元

(1) 项目实施期内，累计 6 人次参加国内学术会议，每次差旅费（含会议注册费）3000 元，共需外埠差旅费： $6 \times 0.3 \text{ 万元/人} = 1.8$ 万元；

(2) 项目组成员往返温氏新兴种猪场和屠宰场采集实验样品 20 人次，每次差旅费 300 元，共需交通费： $20 \times 0.03 \text{ 万元/次} = 0.6$ 万元。

(3) 项目组成员参加 1 人次国际学术会议，包括注册、住宿及来回机票费用等，约需 3.5 万元。

6、出版/文献/信息传播/知识产权事务费：4.2 万元

用于学术论文发表版面费、专利申请费及文献检索费用等。

7、劳务费：7.36 万元

项目执行过程中，每年将有 1 名博士生、2 名硕士生参与研究工作。博士生劳务费 $800 \text{ 元/月} \times 4 \text{ 年} \times 8 \text{ 月} = 2.56$ 万元、硕士生的劳务费 $2 \text{ 人} \times 600 \text{ 元/月} \times 4 \text{ 年} \times 10 \text{ 月} = 4.8$ 万元；共需博硕士劳务费 7.36 万元。

8、专家咨询费：0.41 万元

用于支付来访专家的项目咨询费用。

9、其他支出：无

项目负责人签字：杨杰





国家自然科学基金资助项目签批审核表

<p>我接受国家自然科学基金的资助，将按照申请书、项目批准意见和计划书负责实施本项目（批准号：31972540），严格遵守国家自然科学基金委员会关于资助项目管理、财务等各项规定，切实保证研究工作时间，认真开展研究工作，按时报送有关材料，及时报告重大情况变动，对资助项目发表的论著和取得的研究成果按规定进行标注。</p> <p>项目负责人（签章）：<u>杨杰</u> 2019年9月5日</p>	<p>我单位同意承担上述国家自然科学基金项目，将保证项目负责人及其研究队伍的稳定和研究项目实施所需的条件，严格遵守国家自然科学基金委员会有关资助项目管理、财务等各项规定，并督促实施。</p> <p style="text-align: center;"> 单位（公章） 年 月 日</p>														
本栏目由基金委填写	<p>科学处审查意见：</p> <p style="text-align: center;">2019年9月5日</p> <p>建议年度拨款计划（本栏目为自动生成，单位：万元）：</p> <table border="1" data-bbox="263 996 1013 1108"><tr><th>年度</th><th>总额</th><th>第一年</th><th>第二年</th><th>第三年</th><th>第四年</th><th>第五年</th></tr><tr><td>金额</td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table> <p>负责人（签章）： 2019年11月1日</p>	年度	总额	第一年	第二年	第三年	第四年	第五年	金额						
	年度	总额	第一年	第二年	第三年	第四年	第五年								
金额															
本栏目主要用于重大项目等	<p>科学部审查意见：</p> <p style="text-align: center;">同意科学处审查意见</p> <p>负责人（签章）： 2019年11月8日</p>														
	<p>相关局室审核意见：</p> <p style="text-align: right;">负责人（签章）： 年 月 日</p> <p>委领导审批意见：</p> <p style="text-align: right;">委领导（签章）： 年 月 日</p>														

关于国家自然科学基金资助项目批准及有关事项的通知

顾婷 先生/女士:

根据《国家自然科学基金条例》的规定和专家评审意见,国家自然科学基金委员会(以下简称自然科学基金委)决定批准资助您的申请项目。项目批准号:

31802036, 项目名称: 组蛋白H3K27me3对猪胚胎骨骼肌发育调控作用的研究, 直接费用: 24.00万元, 项目起止年月: 2019年01月至 2021年12月, 有关项目的评审意见及修改意见附后。

请尽早登录科学基金网络信息系统(<https://isisn.nsfc.gov.cn>), 获取《国家自然科学基金资助项目计划书》(以下简称计划书)并按要求填写。对于有修改意见的项目, 请按修改意见及时调整计划书相关内容; 如对修改意见有异议, 须在计划书电子版报送截止日期前提出。

计划书电子版通过科学基金网络信息系统(<https://isisn.nsfc.gov.cn>)上传, 由依托单位审核后提交至自然科学基金委进行审核。审核未通过者, 返回修改后再行提交; 审核通过者, 打印为计划书纸质版(一式两份, 双面打印), 由依托单位审核并加盖单位公章后报送至自然科学基金委项目材料接收工作组。计划书电子版和纸质版内容应当保证一致。向自然科学基金委提交和报送计划书截止时间节点如下:

- 1、提交计划书电子版截止时间为2018年9月11日16点(视为计划书正式提交时间);
- 2、提交计划书电子修改版截止时间为2018年9月18日16点;
- 3、报送计划书纸质版截止时间为2018年9月26日16点。

请按照以上规定及时提交计划书电子版, 并报送计划书纸质版, 未说明理由且逾期不报计划书者, 视为自动放弃接受资助。

附件: 项目评审意见及修改意见表

国家自然科学基金委员会
生命科学部
2018年8月16日



项目批准号	31802036
申请代码	C170102
归口管理部门	
依托单位代码	51064208A0499-0932



国家自然科学基金委员会 资助项目计划书

资助类别：青年科学基金项目

亚类说明：

附注说明：

项目名称：组蛋白H3K27me3对猪胚胎骨骼肌发育调控作用的研究

直接费用：24万元 执行年限：2019.01-2021.12

负责人：顾婷

通讯地址：天河区五山街华南农业大学动科学院616

邮政编码：510642 电 话：020-85283780

电子邮件：tinggu@scau.edu.cn

依托单位：华南农业大学

联系人：唐家林 电 话：020-85280070

填表日期：2018年08月20日

国家自然科学基金委员会制



简表

申请者信息	姓 名	顾婷	性 别	女	出生年月	1986年11月	民 族	汉族
	学 位	博士			职称	讲师		
	是否在站博士后	否			电子邮件	tinggu@scau.edu.cn		
	电 话	020-85283780			个人网页			
	工 作 单 位	华南农业大学						
	所 在 院 系 所	动物科学学院						
依托单位信息	名 称	华南农业大学					代码	51064208A0499
	联 系 人	唐家林			电子邮件	kyc.jhk@scau.edu.cn		
	电 话	020-85280070			网站地址	http://web.scau.edu.cn/kjc/		
合作单位信息	单 位 名 称							
项目基本信息	项 目 名 称	组蛋白H3K27me3对猪胚胎骨骼肌发育调控作用的研究						
	资 助 类 别	青年科学基金项目				亚 类 说 明		
	附 注 说 明							
	申 请 代 码	C170102:家畜遗传育种学				H0417:胚胎着床及早期胚胎发育异常		
	基 地 类 别	国家生猪种业工程技术研究中心						
	执 行 年 限	2019.01-2021.12						
	直 接 费 用	24万元						



项目摘要

中文摘要:

组蛋白甲基化在骨骼肌发育过程中发挥重要调控作用。在啮齿动物中的研究发现,调控骨骼肌发育的主要组蛋白修饰为H3K27me3,但在猪中其作用机制尚不清楚。本项目拟在前期收集杜洛克猪胚胎33、65和90天背最长肌的基础上,采用染色质免疫共沉淀-测序技术(NChIP-Seq)检测组蛋白H3K27me3修饰的DNA序列,绘制猪胚胎33、65和90天背最长肌受H3K27me3调控的首张全基因组图谱,筛选差异修饰的DNA元件和基因;同时采用RNA-Seq技术研究3个时期的转录组,筛选差异表达的基因;在此基础上进行联合分析,初步筛选出关键的DNA调控元件和基因;最后在骨骼肌卫星细胞中对部分关键DNA调控元件和基因进行验证。研究结果将从理论上初步阐明组蛋白H3K27me3对猪骨骼肌发育的调控作用,同时为猪产肉性状的分子改良提供新思路。

Abstract:

Histone methylation plays important roles in regulating muscle development. Studies in mice indicated that H3K27me3 is involved in regulating proliferation and differentiation of skeletal muscle, but the mechanism regulating porcine muscle development is unclear. In the present proposal, we are planning to firstly using native ChIP-sequencing (NChIP-Seq) technique to study the genomic sequences regulated by H3K27me3 in the longissimus dorsi muscle from Duroc embryos on gestation days 33, 65 and 90 which we have already collected before; secondary, using RNA-Seq to study the transcriptome and differentiated expressed genes in skeletal muscle of these three time periods and then select the key DNA sequences and genes combining the data from these two results; finally we will confirm the function of genes and regulated DNA regulatory elements in microsatellite cells. Our project will uncover the function of H3K27me3 in regulating muscle development, which will also provide new idea for molecular breeding to improve pork production.

关键词(用分号分开): 猪; 表达调控; H3K27me3; 骨骼肌; 全基因组

Keywords(用分号分开): pig; expression regulation; H3K27me3; skeletal muscle; whole genome



项目组主要成员

编号	姓名	出生年月	性别	职称	学位	单位名称	电话	证件号码	项目分工	每年工作时间(月)
1	顾婷	1986.11	女	讲师	博士	华南农业大学	020-85283780	420101198611287028	项目负责人	8
2	刘满清	1977.09	女	高级实验师	硕士	华南农业大学	020-85280277	320102197709204620	胚胎培养和移植	6
3	郑恩琴	1981.08	女	实验师	硕士	华南农业大学	020-85283780	441322198108016041	克隆胚胎制备	6
4	徐铮	1979.12	男	兽医师	硕士	华南农业大学	13602496031	430626197912280032	细胞培养和分子生物学实验	6
5	敖政	1991.10	男	博士生	学士	华南农业大学	15820232919	522101199110195416	克隆猪采样分析	10
6	全绒	1992.12	女	硕士生	学士	华南农业大学	18344322116	520402199212280049	细胞培养	10
7	甘炎民	1994.08	男	硕士生	学士	华南农业大学	18819467494	445321199408160311	细胞培养和分子生物学操作	10
总人数		高级		中级		初级		博士后	博士生	硕士生
7		1		3					1	2



国家自然科学基金项目直接费用预算表（定额补助）

项目批准号：31802036

项目负责人：顾婷

金额单位：万元

序号	科目名称	金额
1	项目直接费用合计	24.0000
2	1、设备费	0.0000
3	(1)设备购置费	0.0000
4	(2)设备试制费	0.0000
5	(3)设备升级改造与租赁费	0.0000
6	2、材料费	5.5400
7	3、测试化验加工费	12.6000
8	4、燃料动力费	0.0000
9	5、差旅/会议/国际合作与交流费	1.12
10	6、出版/文献/信息传播/知识产权事务费	2.10
11	7、劳务费	2.6400
12	8、专家咨询费	0.0000
13	9、其他支出	0.0000



国家自然科学基金资助项目签批审核表

<p>我接受国家自然科学基金的资助，将按照申请书、项目批准意见和计划书负责实施本项目（批准号：31802036），严格遵守国家自然科学基金委员会关于资助项目管理、财务等各项规定，切实保证研究工作时间，认真开展研究工作，按时报送有关材料，及时报告重大情况变动，对资助项目发表的论著和取得的研究成果按规定进行标注。</p> <p>项目负责人（签章）： 年 月 日</p>	<p>我单位同意承担上述国家自然科学基金项目，将保证项目负责人及其研究队伍的稳定和研究项目实施所需的条件，严格遵守国家自然科学基金委员会有关资助项目管理、财务等各项规定，并督促实施。</p> <p>依托单位（公章） 年 月 日</p>														
<p>本栏目由基金委填写</p>	<p>科学处审查意见：</p> <p>建议年度拨款计划（本栏目为自动生成，单位：万元）：</p> <table border="1"> <thead> <tr> <th>年度</th> <th>总额</th> <th>第一年</th> <th>第二年</th> <th>第三年</th> <th>第四年</th> <th>第五年</th> </tr> </thead> <tbody> <tr> <td>金额</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table> <p>负责人（签章）： 年 月 日</p>	年度	总额	第一年	第二年	第三年	第四年	第五年	金额						
	年度	总额	第一年	第二年	第三年	第四年	第五年								
	金额														
<p>科学部审查意见：</p> <p>负责人（签章）： 年 月 日</p>															
<p>本栏目主要用于重大项目等</p>	<p>相关局室审核意见：</p> <p>负责人（签章）： 年 月 日</p>														
	<p>委领导审批意见：</p> <p>委领导（签章）： 年 月 日</p>														

检索证明

根据委托人提供的论文材料，委托人华南农业大学国家生猪种业工程技术研究中心 郑恩琴 14 篇论文收录情况如下表。

序号	论文名称	发表刊物及发表的年月卷期/页码等	作者排名	论文等级	作者文中单位	收录情况	影响因子	中科院大类分区
1	Protein Dynamic Landscape during Mouse Mammary Gland Development from Virgin to Pregnant, Lactating, and Involuting Stages	JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY 出版年: 2024 出版日期: MAR 21 卷期: 72 13 页码: 7546-7557 文献类型: Article	共同通讯作者	T2 类	华南农业大学	SCI	IF2-year=6.2 IF5-year=6.4 (2024)	农林科学 1 区 Top 期刊: 是 (2025)
2	Genome-wide detection of multiple variants associated with teat number in French Yorkshire pigs	BMC GENOMICS 出版年: 2024 出版日期: JUL 25 卷期: 25 1 页码: - 文献号: 722 文献类型: Article	最后通讯作者	A 类	华南农业大学	SCI	IF2-year=3.7 IF5-year=4.2 (2024)	生物学 2 区 Top 期刊: 否 (2025)
3	Increased Accuracy of Genomic Prediction Using Preselected SNPs from GWAS with Imputed Whole-Genome Sequence Data in Pigs	ANIMALS 出版年: 2023 出版日期: DEC 卷期: 13 24 页码: - 文献号: 3871	最后通讯作者	A 类	华南农业大学	SCI	IF2-year=2.7 IF5-year=3.0 (2023)	农林科学 2 区 Top 期刊: 否 (2023)

		文献类型: Article						
4	Identification of Candidate Genes for Economically Important Carcass Cutting in Commercial Pigs through GWAS	ANIMALS 出版年: 2023 出版日期: OCT 卷期: 13 20 页码: - 文献号: 3243 文献类型: Article	最后通讯作者	A 类	华南农业大学	SCI	IF2-year=2.7 IF5-year=3.0 (2023)	农林科学 2 区 Top 期刊: 否 (2023)
5	Identification of candidate genes associated with carcass component weights in commercial crossbred pigs through a combined GWAS approach	JOURNAL OF ANIMAL SCIENCE 出版年: 2023 出版日期: JAN 3 卷期: 101 页码: - 文献号: skad121 文献类型: Article	最后通讯作者	A 类	华南农业大学	SCI	IF2-year=2.7 IF5-year=3.2 (2023)	农林科学 2 区 Top 期刊: 否 (2023)
6	Integrated Single-Trait and Multi-Trait GWASs Reveal the Genetic Architecture of Internal Organ Weight in Pigs	ANIMALS 出版年: 2023 出版日期: MAR 卷期: 13 5 页码: - 文献号: 808 文献类型: Article	最后通讯作者	A 类	华南农业大学	SCI	IF2-year=2.7 IF5-year=3.0 (2023)	农林科学 2 区 Top 期刊: 否 (2023)

7	Treatment of Donor Cells with Oxidative Phosphorylation Inhibitor CPI Enhances Porcine Cloned Embryo Development	ANIMALS 出版年: 2024 出版日期: MAY 卷期: 14 9 页码: - 文献号: 1362 文献类型: Article	共同通讯作者	A 类	华南农业大学	SCI	IF2-year=2.7 IF5-year=3.2 (2024)	农林科学 2 区 Top 期刊: 否 (2025)
8	Genome-Wide Association Studies for Flesh Color and Intramuscular Fat in (Duroc x Landrace x Large White) Crossbred Commercial Pigs	GENES 出版年: 2022 出版日期: NOV 卷期: 13 11 页码: - 文献号: 2131 文献类型: Article	最后通讯作者	B 类	华南农业大学	SCI	IF2-year=3.5 IF5-year=3.9 (2022)	生物学 3 区 Top 期刊: 否 (2022)
9	染色质转座酶可及性测序研究进展	遗传 出版年: 2020 卷期: 页码: - 文献号: 文献类型:	最后通讯作者	B 类	华南农业大学	北大核心	无	无
10	瘦肉型猪基因组中 ROH 的应用及研究进展	中国畜牧杂志 出版年: 2022 卷期: 页码: - 文献号: 文献类型:	最后通讯作者	C 类	华南农业大学	北大核心	无	无

11	不同饲养模式对母猪分娩前后行为影响的研究	家畜生态学报 出版年: 2019 卷期: 页码: - 文献号: 文献类型:	通讯作者	C 类	华南农业大学	北大核心	无	无
12	Gender Control of Mouse Embryos by Activation of TLR7/8 on X Sperm via Ligands dsRNA-40 and dsRNA-DR	MOLECULES 出版年: 2024 出版日期: JAN 卷期: 29 1 页码: - 文献号: 262 文献类型: Article	共同通讯作者	B 类	华南农业大学	SCI	IF2-year=4.6 IF5-year=5.0 (2024)	化学 3 区 Top 期刊: 否 (2025)
13	A Transcriptome Analysis Reveals that Hepatic Glycolysis and Lipid Synthesis Are Negatively Associated with Feed Efficiency in DLY Pigs	SCIENTIFIC REPORTS 出版年: 2020 出版日期: JUN 18 卷期: 10 1 页码: - 文献号: 9874 文献类型: Article	共同通讯作者	B 类	华南农业大学	SCI	IF2-year=4.38 IF5-year=5.134 (2020)	综合性期刊 3 区 Top 期刊: 否 (2020)
14	Brain Transcriptome Analysis Reveals Potential Transcription Factors and Biological Pathways Associated with Feed Efficiency in Commercial DLY Pigs	DNA AND CELL BIOLOGY 出版年: 2021 出版日期: FEB 1 卷期: 40 2 页码: 272-282 文献类型: Article	共同通讯作者	B 类	华南农业大学	SCI	IF2-year=3.55 IF5-year=3.557 (2021)	生物学 4 区 Top 期刊: 否 (2021)

说明: 论文等级和中科院大类分区按《华南农业大学学术论文评价方案(试行)》划分。

报告免责声明: 如未盖章, 报告无效

检索员：刘汉忠
华南农业大学图书馆



华南农业大学图书馆SCAU LIB202519377

Protein Dynamic Landscape during Mouse Mammary Gland Development from Virgin to Pregnant, Lactating, and Involuting Stages

Wenjing Wang,⁺ Shunbo Wang,⁺ Hao Wang, Enqin Zheng,^{*} Zhenfang Wu,^{*} and Zicong Li^{*}Cite This: *J. Agric. Food Chem.* 2024, 72, 7546–7557

Read Online

ACCESS |



Metrics & More



Article Recommendations



Supporting Information

ABSTRACT: The mammary gland undergoes significant physiological changes as it undergoes a transition from virgin to pregnancy, lactation, and involution. However, the dynamic role of proteins in regulating these processes during mouse mammary gland development has not been thoroughly explored. In this study, we collected mouse mammary gland tissues from mature virgins aged 8–10 weeks (V), day 16 of pregnancy (P16d), day 12 of lactation (L12d), day 1 of forced weaning (FW 1d), and day 3 of forced weaning (FW 3d) stages for analysis using DIA-based quantitative proteomics technology. A total of 3,312 proteins were identified, of which 843 were DAPs that were categorized into nine clusters based on their abundance changes across developmental stages. Notably, DAPs in cluster 2, which peaked at the L12d stage, were primarily associated with mammary gland development and lactation. The protein–protein interaction network revealed that the epidermal growth factor (EGF) was central to this cluster. Our study provides a comprehensive overview of the mouse mammary gland development proteome and identifies some important proteins, such as EGF, Janus kinase 1 (JAK1), and signal transducer and activator of transcription 6 (STAT6) that may serve as potential targets for future research to provide guidelines for a deeper understanding of the developmental biology of mammary glands.

KEYWORDS: development, mammary gland, mouse, DIA-based quantitative proteomics, EGF

1. INTRODUCTION

Mammary gland is a complex secretory organ composed of various cell types such as fibroblasts, vascular endothelial cells, fat cells, and epithelial cells. The development of mammary gland occurs mainly after birth.¹ The mammary gland serves the crucial function of producing milk to provide essential nourishment for the survival and growth of offspring, which plays a pivotal role in the reproductive success of all mammals.^{2,3} Abnormal mammary gland development increases the morbidity, mortality, and the risk of mental retardation in progenies,^{4,5} whereas aberrant mammary gland growth will cause failure of breastfeeding, which increases the maternal risk of invasive breast cancer, ovarian cancer and type 2 diabetes mellitus.^{5–7} Therefore, uncovering the molecular mechanisms that regulate mammary gland development and lactogenesis will help to improve livestock production and human health.

With the development of molecular biological technology, high-throughput omics studies have significantly enriched our knowledge of gene expression regulation during mammary gland development. Gene expression regulation can manifest at various levels. At the transcriptional level, numerous studies have delved into mammary gland development. For instance, Xuan et al.⁸ used the transcriptome of goat mammary gland tissue to study the physiological aspects of lactation and the molecular genetic mechanisms regulating mammary gland development. Bach et al.⁹ and Twigger et al.¹⁰ determined the gene expression profile of mammary epithelial cells (MECs) at distinct developmental stages through single-cell RNA sequencing, elucidating the hierarchy within MECs. However,

proteomic studies related to mammary gland development remain relatively underexplored. Davies et al.¹¹ identified 18 proteins utilizing mass spectrometry (MS) analysis in forty-four molecular cluster indexes, revealing altered expression patterns from lactation to involution. Nevertheless, substantial-scale proteomic research on mammary gland development is still conspicuously lacking.

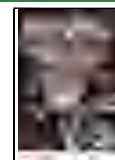
The purpose of this study was to investigate the dynamic proteomic alterations during mammary gland development and to identify potential proteins that may play pivotal roles in regulating mammary gland development at critical time points. The mouse mammary gland tissues were obtained from female ICR mice at five distinct stages, including mature virgin (V), day 16 of pregnancy (P16d), day 12 of lactation (L12d), day 1 of forced weaning (FW 1d), and day 3 of forced weaning (FW 3d). For the first time, this study employed DIA quantification proteomics technology to comprehensively characterize proteins of mouse mammary gland at five distinct stages. The findings of this research expand the existing protein database for murine mammary gland tissue.

Received: December 19, 2023

Revised: March 10, 2024

Accepted: March 12, 2024

Published: March 21, 2024



2. MATERIALS AND METHODS

2.1. Animals. This study received approval from the Ethics Committee of the Laboratory Animal Center at South China Agricultural University (permit number: SYXK-2019–0136). Specific-pathogen-free (SPF) ICR mice, approximately 7 weeks old, were procured (consisting of 15 females and 8 males) from Liaoning Changsheng Biotechnology Co., Ltd. in Benxi, Liaoning, China. These mice were accommodated in an environment with a 12 h light/dark cycle, at a temperature of 22 ± 1 °C, and provided with free access to both food and tap water.

2.2. Mammary Gland Tissue Collection. Fifteen female mice were subjected to random allocation across five groups (with $n = 3$ groups each), including the following categories: V, P16d, L12d, FW 1d, and FW 3d. Notably, except for the mature virgin mice, female mice were paired with males for mating and the presence of vaginal plugs was assessed the morning after pairing. The day on which a vaginal plug was observed marked the initiation of pregnancy (P0d), and the date of delivery signified the commencement of lactation (L1d). Lactating mice were introduced to forced weaning on day L12d, and the day of removing the suckling pups was regarded as day 0 of forced weaning (FW0d). The initial and third days of weaning designated as day 1 of forced weaning (FW 1d, within 24 h) and day 3 of forced weaning (FW 3d, within 72 h). Under pentobarbital sodium anesthesia (45 mg/kg body weight), all female mice were humanely euthanized via cervical dislocation to procure mammary tissue samples. Subsequently, thoracic and abdominal mammary glands were collected following lymph node resection, swiftly frozen with liquid nitrogen, and stored at -80 °C.¹²

2.3. Protein Extraction. The protein extraction procedure was carried out as follows: (1) the appropriate amount of the sample was weighed and placed in a 1.5 mL centrifuge tube. To each tube, a 5 mm steel bead, the necessary quantity of lysis buffer 3, phenyl-methylsulfonyl fluoride (PMSF) with a final concentration of 1 mM, and ethylenediaminetetraacetic acid (EDTA) with a final concentration of 2 mM were added. The contents were vortexed and incubated for 5 min, and then dithiothreitol (DTT) with a final concentration of 10 mM was introduced. (2) A tissue grinder was set at a frequency of 50 Hz for 2 min. Subsequently, the mixture was centrifuged at 25,000g at 4 °C for 20 min and the supernatant was collected. (3) DTT was added with a final concentration of 10 mM, followed by a water bath at 56 °C for 1 h. Once the sample returns to room temperature, iodoacetamide (IAM) with a final concentration of 55 mM was introduced, and the mixture was incubated in a dark environment for 45 min. (4) Cold acetone was added in a volume four times that of the sample, and the tube was placed at -20 °C for 2 h. This step was repeated two to three times until the supernatant becomes colorless. Then, centrifugation was performed at 25,000g, 4 °C for 20 min and the supernatant was discarded. (5) An appropriate amount of Lysis Buffer 3 was added to precipitate, followed by ultrasonication to dissolve the precipitated proteins and centrifugation at 25,000g, 4 °C for 20 min to obtain the supernatant, which was used for quantification, and then stored at -80 °C in a refrigerator.

2.4. Protein Extraction Quality Control and Digestion. Standard proteins (0.2 $\mu\text{g}/\mu\text{L}$ BSA) 0, 2, 4, 6, 8, 10, 12, 14, 16, and 18 μL were sequentially added to the 96-well microtiter plates A1–A10, followed by the addition of 20, 18, 16, 14, 12, 10, 8, 6, 4, and 2 μL of pure water, and then 180 μL of Coomassie Brilliant Blue G-250 Quantitative Working Solution was added to each well. The OD595 was determined using a microplate reader, and a linear standard curve was constructed based on OD595 readings and corresponding protein concentrations. The protein solution to be tested was subjected to multiple dilutions, and 180 μL of the Quantitative Working Solution was added to 20 μL of the diluted protein solution. The OD595 was measured, and the sample protein concentration was determined by referencing the standard curve. The quality of the samples was assessed through SDS-PAGE and Coomassie Brilliant Blue R-250 staining. Protein digestion was performed as follows: (1) 100 μg of protein solution was taken per sample and diluted with 50 mM NH_4HCO_3 in a 4:1 ratio. (2) A total

of 2.5 μg of Trypsin enzyme was added at a protein-to-enzyme ratio of 40:1 and incubated at 37 °C for 4 h. (3) The enzymatic peptides were desalted using a Strata X column and subsequently dried under vacuum.

2.5. High pH RP Separation. Peptides from all samples were pooled in equal amounts and diluted with mobile phase A (5% ACN, pH 9.8). The Shimadzu LC-20AB HPLC system coupled with a Gemini high pH C18 column (5 μm , 4.6×250 mm) was used. The sample was introduced to the column and then eluted at a flow rate of 1 mL/min by gradient: 5% mobile phase B (95% ACN, pH 9.8) for 10 min, a transition from 5 to 35% mobile phase B for 40 min, a further shift to 35 to 95% mobile phase B for 1 min, flow phase B that lasted for 3 min, and finally, equilibration with 5% mobile phase B for 10 min. The elution peak was monitored at a wavelength of 214 nm, and the component was collected every minute. Components were combined into a total of 10 fractions and subsequently subjected to freeze-drying.

2.6. DDA and DIA analysis by Nano-LC-MS/MS. The dried peptide samples were reconstituted with mobile phase A (2% ACN, 0.1% FA), followed by centrifugation at 20,000g for 10 min, and the supernatant was taken for injection. The chromatographic separation was conducted by utilizing a Thermo UltiMate 3000 UHPLC liquid chromatograph. The sample was first enriched in the trap column, desalted, then entered a tandem self-packed C18 column (150 μm internal diameter, 1.8 μm column size, 35 cm column length), and separated at a flow rate of 500 nL/min following the subsequent gradient profile: 0 to 5 min, maintained at 5% mobile phase B (98% ACN, 0.1% FA); 5 to 130 min, mobile phase B linearly increased from 5 to 25%; 130 to 150 min, mobile phase B rose from 25 to 35%; 150 to 160 min, mobile phase B rose from 35 to 80%; 160 to 175 min, maintaining 80% mobile phase B; 175 to 175.5 min, mobile phase B decreased from 80 to 5%; 175.5 to 180 min, held at 5% mobile phase B. The nanoliter liquid phase separation end point was directly connected to the mass spectrometer with the subsequent settings.

For data-dependent acquisition (DDA) analysis, LC-separated peptides were subjected to ionization via nanoESI and introduced into the tandem mass spectrometer, Fusion Lumos (Thermo Fisher Scientific, San Jose, CA) employing DDA mode. The key instrumental parameters were configured as follows: ion source voltage of 2 kV; MS scan range of 350–1500 m/z ; MS resolution of 60,000; maximal injection time (MIT) of 50 ms; MS/MS collision type higher-energy collisional dissociation (HCD), normalized collision energy (NCE) 30; MS/MS resolution of 15,000; MIT of 50 ms; dynamic exclusion duration of 30 s. The commencement m/z for MS/MS was fixed to 100. The selection criteria for the precursor MS/MS scan were as follows: charge range from 2+ to 6+, choosing the top 30 precursors with an intensity exceeding 2E4. AGC was MS 3E6, MS/MS 1E5.

For data-independent acquisition (DIA) analysis, LC-separated peptides were subjected to ionization through nanoESI and introduced into the tandem mass spectrometer Fusion Lumos (Thermo Fisher Scientific, San Jose, CA), operating in the DIA mode. The primary instrument settings were as follows: ion source voltage of 2 kV; MS scan range of 400–1500 m/z ; MS resolution of 60,000; MIT of 50 ms; the range of 400–1500 m/z was evenly partitioned into 44 contiguous windows for MS/MS scanning. For the MS/MS scans, the following parameters were utilized: MS/MS collision type HCD, MIT of 54 ms. Fragment ions were scanned in Orbitrap with an MS/MS resolution of 30,000 and collision energy of 30. Additionally, the AGC was configured to 5E4.

2.7. Data Analysis. The raw data generated from DDA-MS were subjected to identification using the Andromeda search engine¹³ within MaxQuant software (version 1.5.3.30).¹⁴ The search parameters were defined as follows: (1) digestion enzyme: trypsin; (2) minimal peptide length: 7; (3) fixed modifications: carbamidomethyl (C); (4) variable modifications: oxidation (M), acetyl (protein N-term), Gln \rightarrow pyro-Glu (N-term Q), and deamidated (NQ); (5) protein false discovery rate (FDR) and PSM-level FDR: 0.01. The MaxQuant output file was utilized to create the final spectral library. The raw data from DIA-MS was analyzed using the iRT peptides for

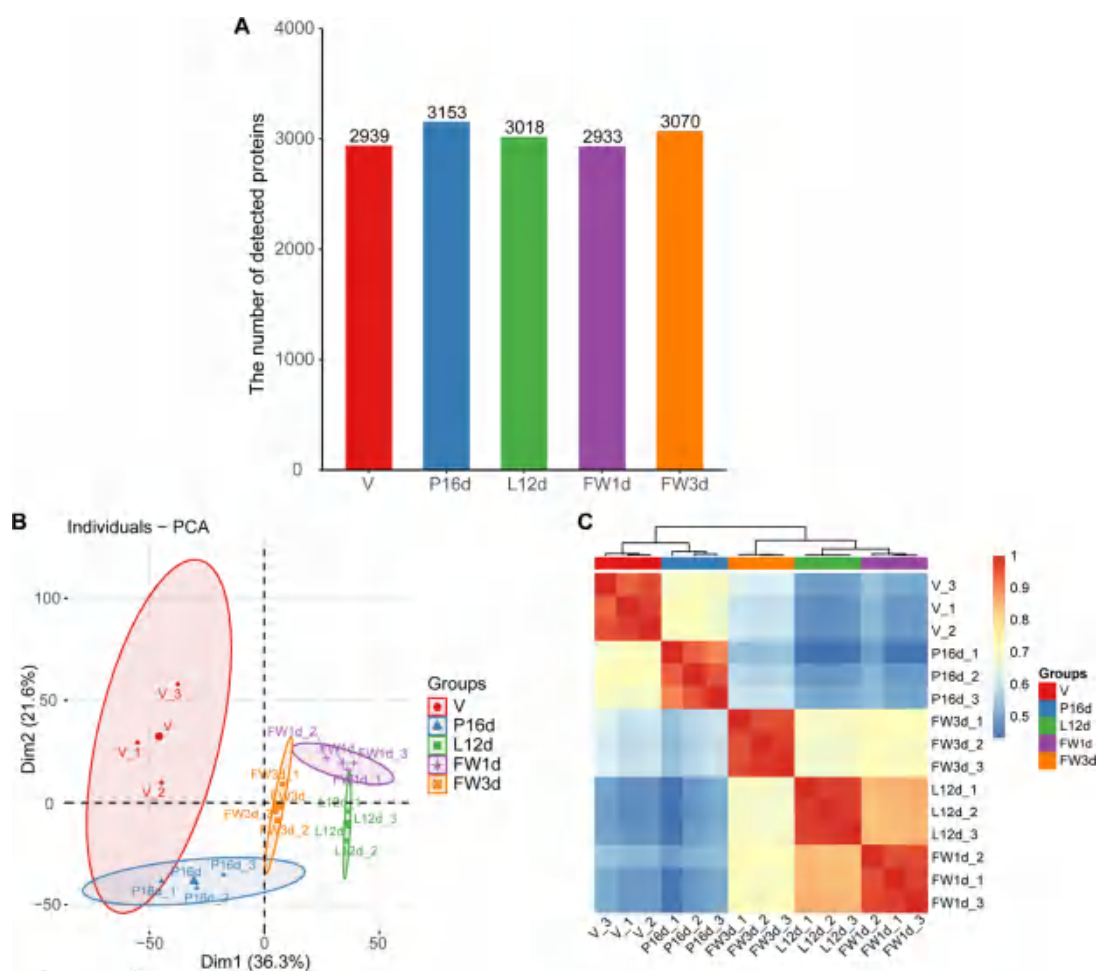


Figure 1. Global information for the proteome during mouse mammary gland development. (A) Number of proteins identified in V, P16d, L12d, FW 1d, and FW 3d stages of mouse mammary gland tissues. (B) Principal component analysis of the total identified proteins. Same color represents biological replicates of the same stage, and the dot located at the center of each circle represents the average protein abundance of each group. (C) Unsupervised hierarchical clustering and heatmap of proteins identified from 15 mouse mammary gland samples.

retention time calibration. Subsequently, leveraging the target-decoy model, which is applicable to SWATH-MS, a false positive control was performed with FDR 1%. This rigorous approach was adopted to yield statistically significant quantitative results.

2.8. Bioinformatic Analysis. The protein abundance data obtained from all five stages were evaluated by ANOVA and Tukey's honest significant difference test. Subsequently, *p*-values were corrected using the Benjamini–Hochberg method to obtain the FDR values. Differentially abundant proteins (DAPs) were identified with the screening criteria of $|\log_2(\text{fold change})| \geq 1$ and $\text{FDR} < 0.05$, signifying statistical significance. The subcellular localization analysis of DAPs was conducted through DeepLoc (version 2.0, accessible at <https://services.healthtech.dtu.dk/service.php?DeepLoc-2.0>). Gene Ontology (GO) term analysis for the DAPs was performed using the clusterProfiler R package (version 4.2.1).¹⁵ To explore common protein abundance patterns across the five stages of mammary gland development, Mfuzz R software (version 2.54.0) was employed.¹⁶ Furthermore, protein–protein interactions (PPI) were investigated using the STRING database (<http://string-db.org>).¹⁷

2.9. Immunofluorescence of Histological Sections. Detailed protocols for immunofluorescent staining of histological sections can be found elsewhere.¹⁸ Briefly, the fixed tissues were embedded, sectioned, deparaffinized, rehydrated, antigen-retrieved, and blocked with serum before being incubated with primary antibodies. The following primary antibodies were used: anti-EGF antibody (catalog number ab288294) from Abcam; cytokeratin 18 polyclonal antibody

(catalog number 10830–1-AP) from Proteintech. The CY3-TSA (red) and FITC-TSA (green) were used for the detection of the specific targets. Slides were counterstained briefly with DAPI to visualize the nuclei. Fluorescence images of the histological sections on the slides were taken on a fluorescence microscope (Nikon, Tokyo, Japan) with the same exposure time and intensity.

3. RESULTS

3.1. Overview of the Quantitative Proteomics Analysis. A total of 3,312 proteins were identified in this study. Specifically, 2,939 proteins were detected in the V stage, 3,153 in the P16d stage, 3,018 in the L12d stage, 2,933 in the FW 1d stage, and 3,070 in the FW 3d stage (Figure 1A, Supporting Information Table S1). Principle component analysis (PCA) revealed that samples from five different stages exhibited clear separation and samples from the same stage could be clustered together well (Figure 1B). Furthermore, the unsupervised hierarchical clustering and heatmap of proteins in the five stages (Figure 1C) underscored a robust correlation among samples within each stage, affirming the reliability of our proteomics data set.

3.2. Attributes of Stage-Specific Proteins Identified during Mammary Gland Development. The 3,312 proteins were detected in at least one of the stages, with 2,673 proteins identified across all stages. Additionally, 16, 72,

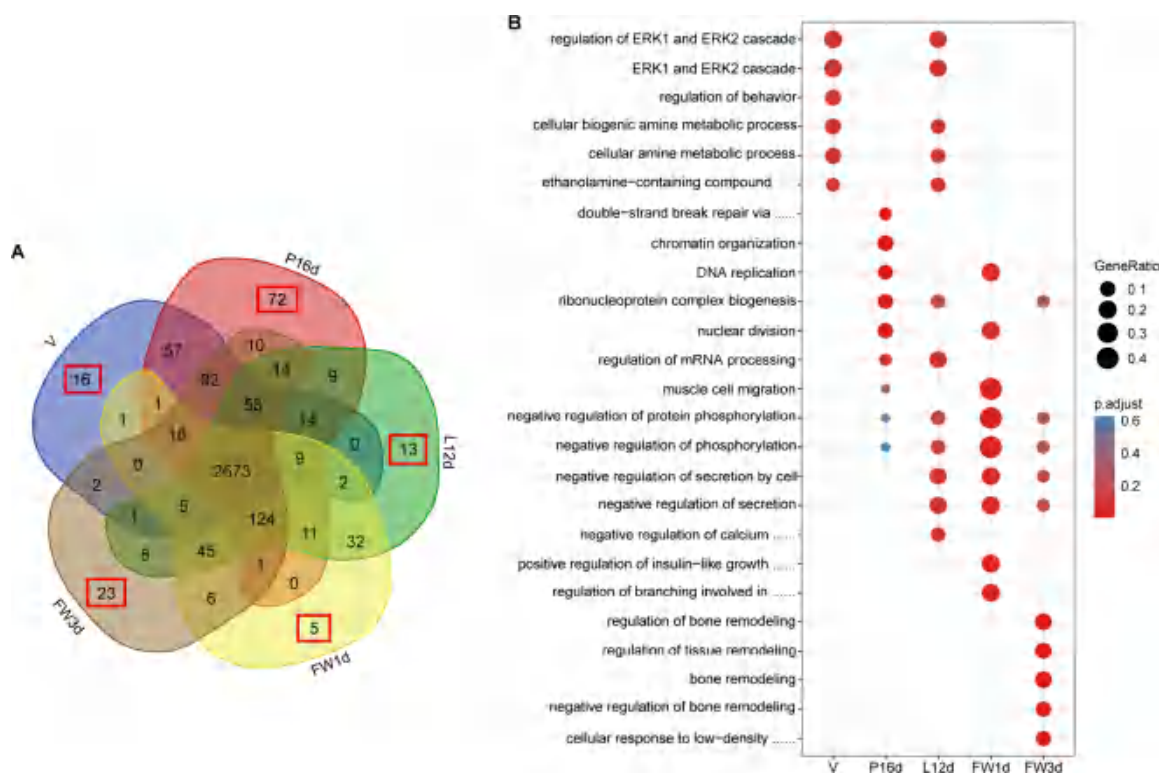


Figure 2. Functional analysis of stage-specific proteins. (A) Overlap of proteins identified in mouse mammary gland at each stage. (B) Functional analysis of the stage-exclusive proteins. The size of the circle represents the ratio of proteins enriched in the GO term and red represents more significant enrichment.

13, 5, and 23 proteins were exclusively found in the V, P16d, L12d, FW 1d, and FW 3d stages, respectively (Figure 2A, Supporting Information, Table S2). Subsequently, these stage-specific proteins were annotated through GO analysis. Specifically, proteins exclusive to the V stage exhibited enrichment in regulation of ERK1 and ERK2 cascade (GO:0070372) and ERK1 and ERK2 cascade (GO:0070371). The exclusive proteins at P16d stage enriched in DNA replication (GO:0006260) and ribonucleoprotein complex biogenesis (GO:0022613). The exclusive proteins at L12d stage enriched in regulation of mRNA processing (GO:0050684), negative regulation of secretion by cell (GO:1903531), and negative regulation of secretion (GO:0051048). The exclusive proteins at FW 1d stage enriched in negative regulation of protein phosphorylation (GO:0001933) and negative regulation of phosphorylation (GO:0042326). Lastly, the exclusive proteins at the FW 3d stage were enriched in regulation of tissue remodeling (GO:0034103) (Figure 2B).

3.3. Identification of DAPs during Mouse Mammary Gland Development. A total of 843 proteins exhibited significant differential abundance in this study ($FDR < 0.05$, $|\log_2(\text{fold change})| \geq 1$, Table S3). In pairwise comparisons across experimental conditions, comparisons of P16d_V, L12d_P16d, FW 1d_L12d, and FW 3d_FW 1d found that protein abundance initially increases, then decreases, and then increases again as the mammary gland develops. Moreover, the majority of proteins exhibited increased abundance in the P16d, L12d, FW 1d, and FW 3d stages compared to the V stage. Notably, a greater number of proteins at the L12d, FW 1d, and FW 3d stages displayed a gradual decrease in abundance when compared with the P16d stage. The most

prominent changes in the proteome were observed within the L12d_V group, whereas the smallest proteome changes were found in the FW 1d_L12d group. These changes may be related to the proliferation of MECs, branching morphogenesis, alveologenesis, and tissue remodeling during mammary gland development (Figure 3A). However, it is noteworthy that the $\log_2(\text{fold change})$ values of all DAPs identified between the FW 1d and L12d stages (FW 1d_L12d group) were greater than 10. This suggests that from L12d to FW 1d stage, DAPs between these two stages have a big change in their abundance (Figure 3B).

3.4. Cluster Analysis of DAPs during Mouse Mammary Gland Development. To explore the dynamic changes in the abundance of all DAPs across the five consecutive developmental stages, we conducted a time-series cluster analysis using the Mfuzz R package. This analysis partitioned all DAPs into nine distinct clusters, each characterized by unique expression patterns (Figure 4A, Table S4). Among the nine clusters, DAPs in clusters 2, 4, 8, and 9 exhibited peak abundance levels at the L12d stage. The specific protein counts for each cluster were visualized in Figure 4B. Based on GO analysis, each cluster displayed distinct characteristics related to biological processes, molecular functions, and cellular components. For instance, GO terms associated with Golgi vesicle transport were enriched in cluster 4, while those related to ribonucleoprotein complex biogenesis were enriched in cluster 8. Cluster 9 exhibited an enrichment in GO terms associated with ribosome synthesis. Importantly, GO terms pertaining to mammary gland development and lactation were found to be enriched in cluster 2 (Figure 4C).

3.5. PPI Network Analysis of DAPs in Cluster 2. Milk proteins such as kappa-casein (CSN3), alpha-S1-casein

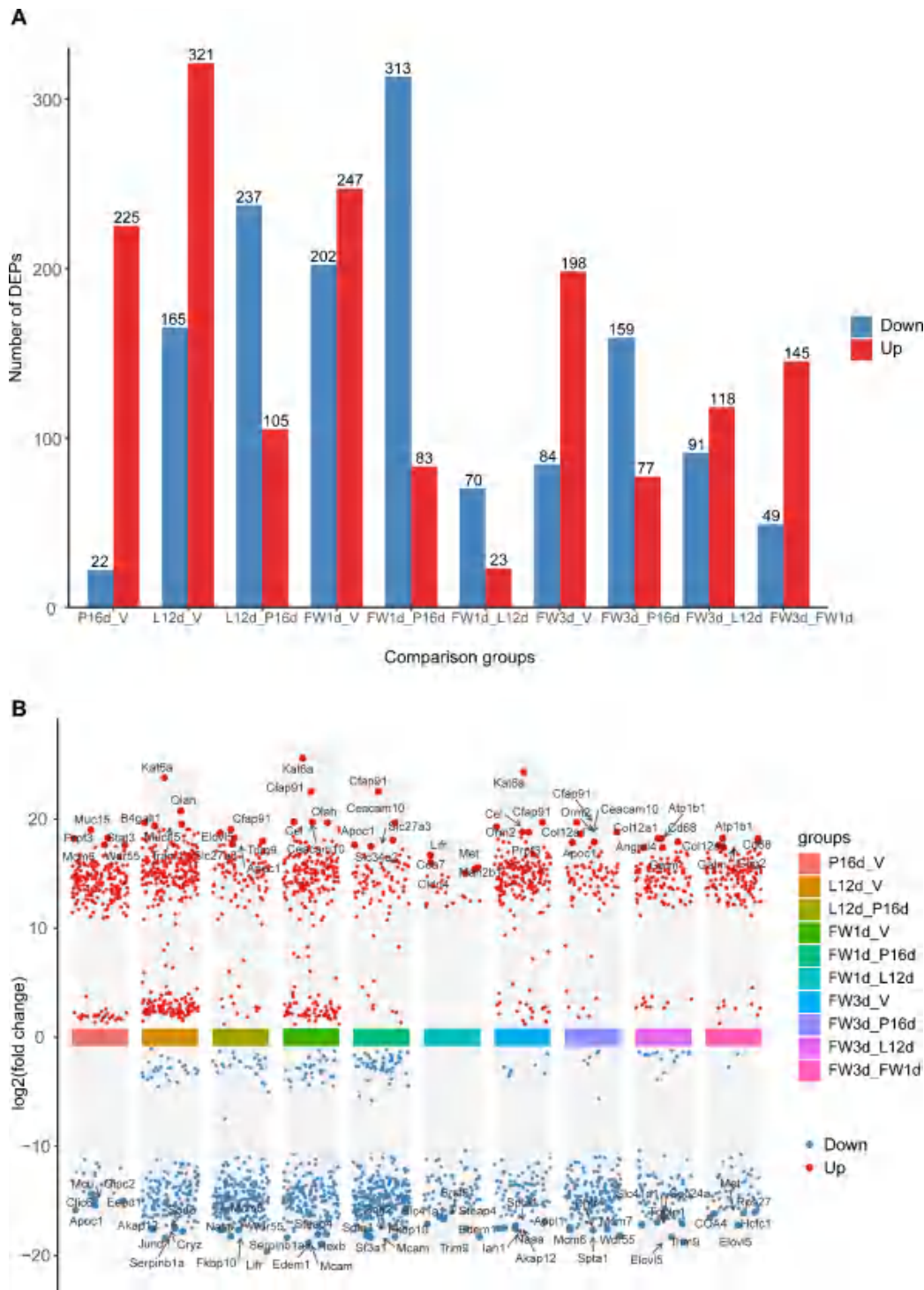


Figure 3. Global information for DAPs during mouse mammary gland development. (A) DAPs with increased and decreased abundance in sequential pairwise comparison of five mouse mammary gland development stages. (B) Log₂ transform of abundances plotted against pairwise comparisons. Only proteins passing the applied threshold ($|\log_2(\text{fold change})| \geq 1$, $\text{FDR} < 0.05$) displayed.

(CSN1S1), alpha-S2-casein-like B (CSN1S2B), and whey acidic protein (WAP) were notably enriched in cluster 2, suggesting a potential association of proteins within this cluster with milk synthesis and secretion processes. Furthermore, comprehending the molecular mechanisms of mammary gland development from a systemic perspective is aided by protein–protein interaction (PPI) analysis. To unveil the protein regulatory networks within the differentially abundant proteins of cluster 2, a PPI network analysis was conducted by using the

STRING database and visualized through Cytoscape software. Within cluster 2, of the 83 proteins, 31 displayed interconnections, with some milk proteins exhibiting high connectivity. Remarkably, EGF emerged as the predominant hub protein within cluster 2, showcasing the highest number of connections. It exhibited close associations with milk proteins while also demonstrating robust connections with JAK1 and STAT6 proteins. The JAK/STAT signaling pathway plays an

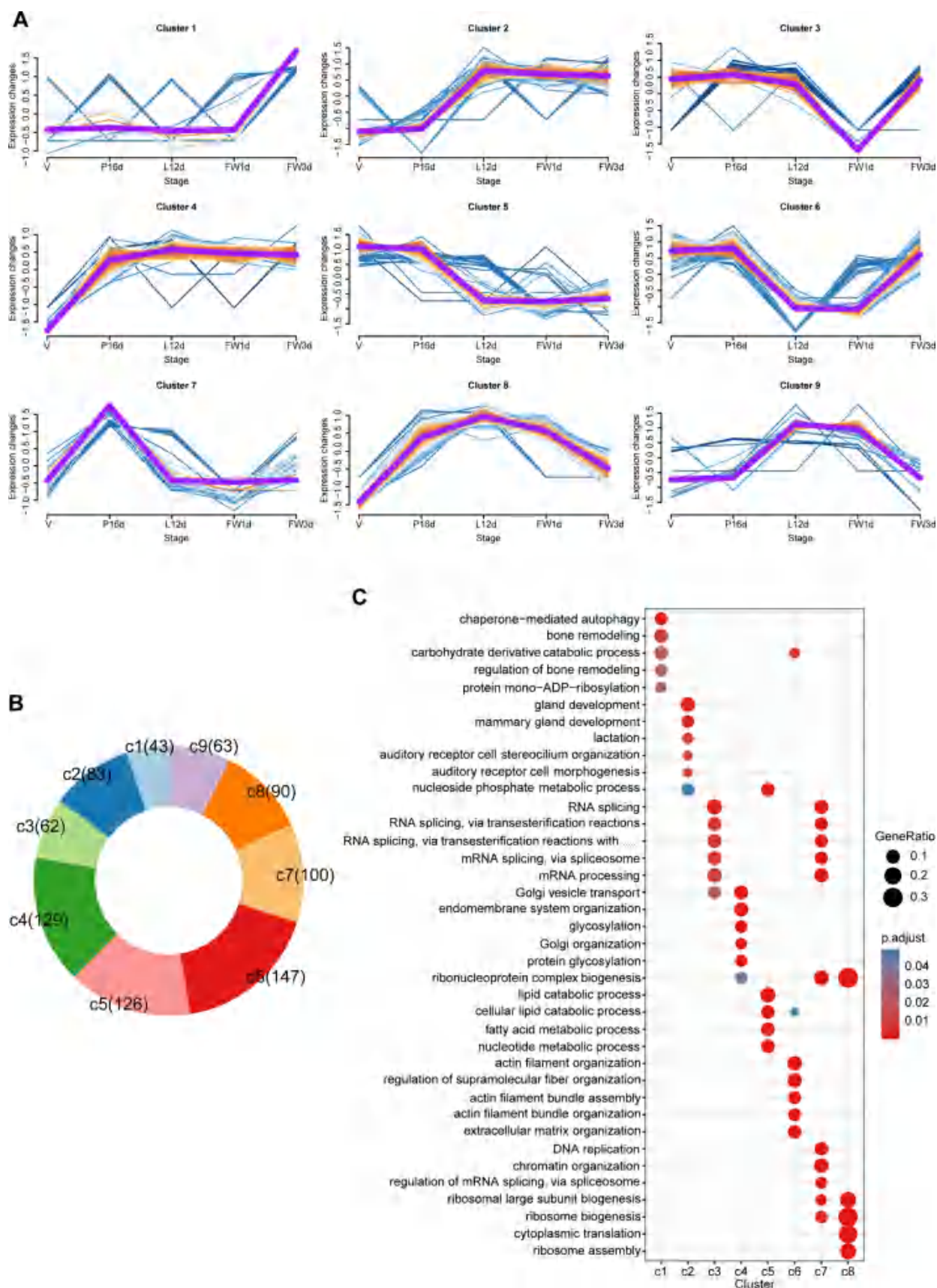


Figure 4. Dynamics of DAPs during mouse mammary gland development. (A) Fuzzy c-means clustering identified nine distinct temporal patterns of proteins expression. The x axis represents five developmental stages, while the y axis represents the \log_2 transform, normalized intensity ratios in each stage. (B) The pie chart shows the number of proteins in each cluster. (C) Functional analysis of protein clusters. The size of the circle represents the ratio of DAPs enriched in the GO term to all proteins in this GO term and red represents more significant enrichment.

important role in the processes of mammary gland development and lactation^{19,20} (Figure 5).

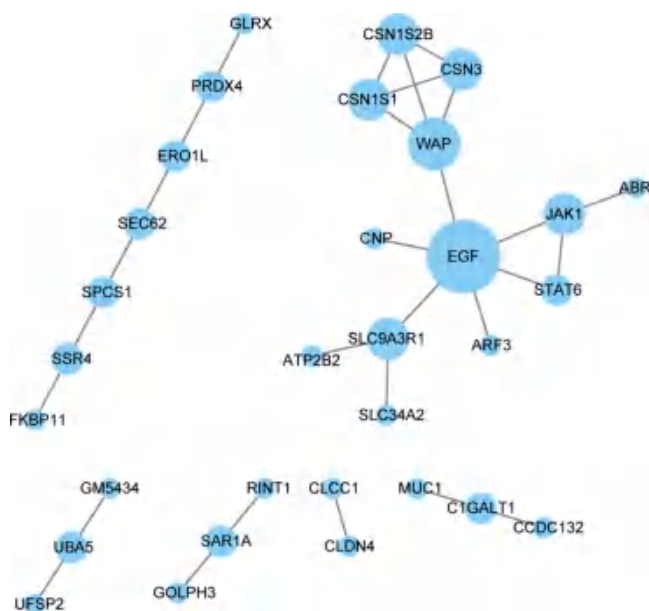


Figure 5. PPI network of the DAPs in cluster 2.

3.6. Immunofluorescent Analysis the Dynamic Abundance of EGF Protein during Mouse Mammary Gland Development.

Analysis of the EGF protein abundance

revealed significant variations across different stages of mammary development. A notable upregulation of EGF was observed at the onset of lactation, in stark contrast to the minimal levels detected in the early stages of mammary gland development, as indicated by immunofluorescence analysis. These results exhibited consistent abundance patterns of EGF with those obtained from DIA analysis, reinforcing the reliability of DIA technology (Figure S1). Additionally, in the mouse mammary gland, luminal epithelial cells have been demonstrated to express cytokeratin proteins 18 (CK18).^{1,21} The protein abundance of EGF was detected in luminal cells (Figure 6).

4. DISCUSSION

We successfully obtained proteomes representing the five stages of mouse mammary development, identifying a total of 3,312 distinct proteins. Notably, our study revealed that the number of identified proteins did not change significantly across stages. The overall protein count remained relatively stable throughout mammary gland development, with a remarkable 88.7% of the identified proteins already present during the early development phase (V stage). Only 3.9% of the identified proteins were unique to the V, P16, L12, FW 1d, and FW 3d stages, totaling 129 corresponding proteins. This finding highlights the persistence of a core set of proteins throughout mammary gland development, whereas a subset of stage-exclusive proteins plays a vital role at specific developmental stages.

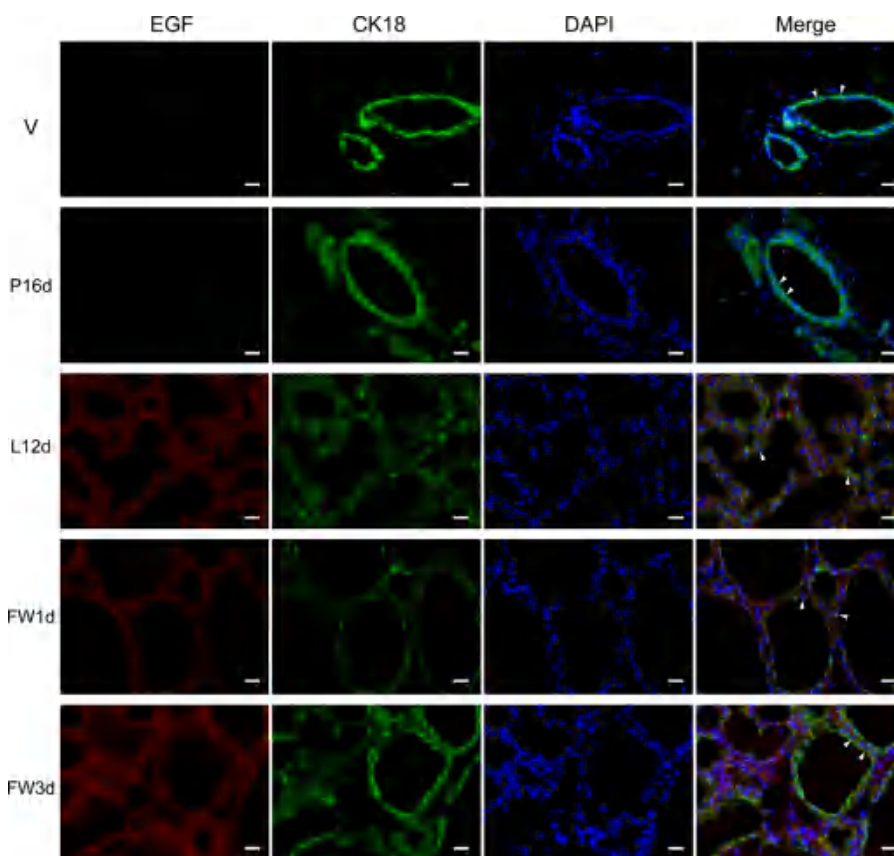


Figure 6. Immunofluorescent staining for EGF protein during mammary gland development. The luminal cell is shown by the arrow without a tail. EGF (red), CK18 (green), DAPI (blue). Scale bar, 20 μ m.

Notably, our functional analysis illuminated the importance of these stage-exclusive proteins in driving specific biological processes. Exclusive proteins at the V stage were enriched in GO terms related to the ERK1 and ERK2 cascade, which were known to regulate essential biological functions, including cellular proliferation and survival.^{22,23} In the gland of virgin mice, ductal complexity increases through the epithelium proliferates and the increase of lateral branches during each estrous cycle.²⁴ Additionally, these enriched GO terms were also observed at the L12d stage, which were known to regulate milk protein synthesis.²⁵ At the P16d stage, exclusive proteins were associated with GO terms related to DNA replication and ribonucleoprotein complex biogenesis, while those at the L12d stage were linked to cell secretion processes. At around midpregnancy, cell differentiation becomes dominant as the gland moves into the secretory initiation phase.²⁶ The alveolar buds progressively cleave and differentiate into individual alveoli, which ultimately become milk-secreting lobules during lactation.²⁷ During lactation, the significant increase in milk synthesis and genes involved in protein synthesis promote secretory activation of alveolar cells. The milk secretion copious into the alveolar lumen.²⁸ Milk is then discharged through the contraction of myoepithelial cells and transported to the nipple through the lactiferous ducts.³ Similarly, proteins exclusive at the FW 1d stage were enriched in GO terms associated with protein phosphorylation, and exclusive proteins at the FW 3d stage were connected to tissue remodeling processes. The involution process is divided into two phases, characterized by changes in gene expression patterns, biochemical reactions, and histological features.²⁹ The first phase is characterized by milk stasis and no major morphological changes. The main apoptotic pathway in this phase is the lysosomal-mediated cell death (LCD) pathway.³⁰ STAT3, a member of the Stat protein family, plays a pivotal role in the involution process, becoming phosphorylated to its active form in the initial phase of involution.^{31–34} The second irreversible phase, occurring after 48 h of weaning, is marked by the predominance of mitochondria-mediated intrinsic apoptotic pathways, facilitating mammary gland remodeling.³⁵ Interestingly, we found that GO terms associated with cell secretion processes were enriched in both the FW 1d and FW 3d stages. STAT3 was activated by secreted cytokines of MECs, such as leukemia inhibitory factor (LIF), interleukin-6 (IL6), and transforming growth factor (TGF) β 3 in response to milk stasis.^{34,36,37} In addition, the growth and differentiation factors secreted by extracellular matrix (ECM) contribute to adipogenesis in the second phase of involution.³⁸ Therefore, these findings align with the known biological processes of the mammary gland development.

Although the mammary gland undergoes dynamic developmental changes, the primary proliferation of MECs predominantly occurs during pregnancy.³⁹ At the P16d stage, we identified the highest number of stage-specific proteins, totaling 72 specific proteins. Notably, this group includes proteins such as WDR55, MCM6, MCM7, and DHFR, which are high-abundance proteins involved in the cell proliferation process.^{40,41} Remarkably, this study is consistent with the findings of Davies et al.,¹¹ where the protein of the minichromosome maintenance protein 3 (MCM3) was shown to increase during pregnancy and involution, followed by a decrease during lactation.

A total of 843 differentially abundant proteins were identified in our mammary gland proteome results. Upon

pairwise comparison of protein levels across different developmental stages, an interesting phenomenon emerged, with protein abundance first increasing, then decreasing, and finally increasing again as mammary gland development progressed. This dynamic shift can be attributed to several key factors driving mammary tissue transformation: the proliferation of MECs, ductal branching, rapid expansion of luminal epithelial cells, and the formation of alveoli during pregnancy.^{42,43} Subsequently, during lactation, the primary focus was on maintaining mammary gland milk secretion.³ Postweaning, the milk in the mammary epithelium stagnated, initiating the involution process.⁴⁴ During the FW 1d period, the initial process of involution is reversible, and resumption of suckling reestablishes milk supply.³⁴ As mammary gland involution progresses, the number of differentially abundant proteins gradually increases, particularly in the FW 3d stage of the second irreversible phase. This can be attributed to a significant upregulation of proteases, including matrix metalloproteinases like gelatinase A and stromelysin-1, which regulate mammary gland remodeling and cell apoptosis during this phase.⁴⁵ Among these five periods, the highest number of differentially abundant proteins was observed in the L12d_V group. Notably, B4GALT1, one of the identified proteins in the group, holds significance as a candidate gene affecting milk performance traits and encodes the catalytic component of lactose synthesis.⁴⁶ The smallest number of differentially abundant proteins was observed in the FW 1d_L12d group. Among the proteins identified in this group, STEAP4, a down-regulated protein, has been reported to play a role in regulating adipogenesis.⁴⁷ However, stromal adipocytes undergo significant decreases during lactation.^{48,49} The subsequent weaning phase triggers mammary gland involution, leading to the regression of alveolar cells and rapid expansion of adipocytes to repopulate the stromal tissue.^{50,51}

All of the differentially abundant proteins identified in this study were categorized into nine distinct clusters, suggesting their diverse roles in regulating mammary gland development. This diversity was further validated through a GO enrichment analysis. Of particular note was that only proteins in cluster 2 were enriched in pathways related to mammary gland development and lactation. These milk proteins of WAP, CSN1S1, CSN1S2B, and CSN3 were enriched in cluster 2. Furthermore, EGF was highly enriched in cluster 2, which significantly promoted alveolar development and milk proteins expression.⁵² It is worth highlighting that not only did proteins in cluster 2 reach their peak abundance at the L12d stage, but proteins in clusters 4 and 8 also exhibited peak expression during this stage. The process of lactation is a distinctive physiological feature in mammals, necessitating alveologenesis and the differentiation of alveoli before parturition to synthesize and secrete milk components.²⁷ Specifically, proteins in cluster 4 were primarily associated with protein transport functions. The family of transmembrane emp24 domain containing (TMED) proteins has become key protein transport-regulatory factors.⁵³ TMED3 in cluster 4 can recognize endoplasmic reticulum core-glycosylated membrane protein cargo, and direct them to the cell surface.⁵⁴ Meanwhile, proteins in cluster 8 were predominantly linked to ribosome formation, which is the site of protein synthesis. The mTOR is a key regulator in cellular protein synthesis. Once activated, mTOR complex 1 (mTORC1) is phosphorylated and activates the ribosomal protein S6 kinase 1 (S6K1) to phosphorylate ribosomal protein S6 (RPS6) in cluster 8 to regulate milk

protein synthesis.^{55,56} These findings align with the heightened synthesis and secretion of milk proteins during lactation.⁵⁷ During secretory activation, there was an increase in the content of the endoplasmic reticulum and Golgi within the alveolar cells.⁵⁸ Moreover, a study in mice have demonstrated an increase in the expression of genes that promote milk synthesis during secretory activation.⁵⁹

A protein–protein interaction network was constructed to further investigate the regulation of mammary gland development and lactation. Notably, EGF exhibited interactions with milk proteins and most proteins in cluster 2 within the PPI network. Milk protein is an important component of milk.⁶⁰ Milk proteins are usually divided into two categories: whey proteins and caseins, which are considered the markers of mammary gland secretion differentiation.⁶¹ CSN1S1 accounts for the majority of casein in milk, while CSN3 plays a crucial role in maintaining the structure and stability of casein micelles and affects milk coagulation and cheese quality.⁶² Therefore, the EGF was selected for subsequent analysis. The Mcl-1 protein has been observed to be upregulated in alveolar cells at the switch to lactation, and its deficiency has been shown to impair lactation.⁶³ Furthermore, EGF was recognized as one of the most highly upregulated genes in lactogenesis. Inhibiting EGF significantly hinders lactation and coincides with a reduction in Mcl-1. EGF serves as a pivotal trigger for Mcl-1 translation, ensuring the survival of alveolar cells during lactation.⁶³ In this study, the increased abundance of EGF protein at the L12d stage was confirmed through immunofluorescence analysis, which aligns with findings from prior studies.⁶³ Furthermore, this further suggests that EGF seems to be primarily induced at the onset of lactation. As a member of the JAK family, JAK1 exhibited a widespread expression profile.⁶⁴ The absence of JAK1 leads to embryonic or neonatal lethality.⁶⁵ JAK1 plays a critical role in the cytokine-induced tyrosine phosphorylation of STAT1, STAT3, and STAT6 in mammary epithelial cells. JAK1 mainly plays a discrete role in the activation of STAT proteins in mammary gland during lactation to involution switching, which contributes to the remodeling of the epithelial compartment during postlactational involution.¹⁸ Furthermore, STAT6 also participates in the regulation of alveolar cell development by regulating immune response.⁶⁶ STAT6 is an important regulator for maintaining the balance of lineages in mammary gland.⁶⁷ A previous study showed that *Jak1* and *Stat6* genes were both expressed in luminal epithelial cells.⁹ Therefore, there may be an interaction between EGF, JAK1 and STAT6 proteins to jointly regulate mammary gland development.

In summary, this study employed a DIA-based proteomics approach to reveal proteins associated with specific stages of mammary gland development, identifying 843 DAPs in mouse mammary glands across five crucial developmental stages. To validate the general significance of these findings, we subsequently conducted immunohistochemistry experiments. This research underscores the pivotal roles of EGF, JAK1, STAT6, and a series of milk proteins in mammary gland development and lactation. These findings provide new perspectives on the proteome profile and the dynamic expression of proteins throughout the mammary gland development process, contributing to a deeper understanding of the intrinsic physiological functions of mammary tissue.

■ ASSOCIATED CONTENT

Data Availability Statement

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the iProX partner repository^{68,69} with the data set identifier PXD046407.

■ Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jafc.3c09647>.

(Figure S1) Protein abundance of EGF during mouse mammary gland development through DIA analysis (PDF)

(Table S1) Dynamic abundance of proteins during mammary gland development in mice (XLSX)

(Table S2) Proteins at different stages of mammary gland development in mice (XLSX)

(Table S3) Analysis results of protein abundance difference at different stages of mammary gland development in mice (XLSX)

(Table S4) Differential abundance protein clustering results during mammary gland development in mice (XLSX)

■ AUTHOR INFORMATION

Corresponding Authors

Enqin Zheng – National Engineering Research Center for Breeding Swine Industry, Department of Animal Genetics, Breeding and Reproduction, College of Animal Science, Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, State Key Laboratory Swine and Poultry Breeding Industry, and National and local joint Engineering Research Center for Livestock and Poultry Breeding Industry, South China Agricultural University, Guangzhou 510642, China; Email: eqzheng@scau.edu.cn

Zhenfang Wu – National Engineering Research Center for Breeding Swine Industry, Department of Animal Genetics, Breeding and Reproduction, College of Animal Science, Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, State Key Laboratory Swine and Poultry Breeding Industry, and National and local joint Engineering Research Center for Livestock and Poultry Breeding Industry, South China Agricultural University, Guangzhou 510642, China; Guangdong Provincial Laboratory of Lingnan Modern Agricultural Science and Technology, Guangzhou 510642, China; Email: wzf@scau.edu.cn

Zicong Li – National Engineering Research Center for Breeding Swine Industry, Department of Animal Genetics, Breeding and Reproduction, College of Animal Science, Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, State Key Laboratory Swine and Poultry Breeding Industry, and National and local joint Engineering Research Center for Livestock and Poultry Breeding Industry, South China Agricultural University, Guangzhou 510642, China; Guangdong Provincial Laboratory of Lingnan Modern Agricultural Science and Technology, Guangzhou 510642, China; Email: lizicong@scau.edu.cn

Authors

Wenjing Wang – National Engineering Research Center for Breeding Swine Industry, Department of Animal Genetics,

Breeding and Reproduction, College of Animal Science, Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, State Key Laboratory Swine and Poultry Breeding Industry, and National and local joint Engineering Research Center for Livestock and Poultry Breeding Industry, South China Agricultural University, Guangzhou 510642, China; orcid.org/0000-0002-1726-7815

Shunbo Wang – National Engineering Research Center for Breeding Swine Industry, Department of Animal Genetics, Breeding and Reproduction, College of Animal Science, Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, State Key Laboratory Swine and Poultry Breeding Industry, and National and local joint Engineering Research Center for Livestock and Poultry Breeding Industry, South China Agricultural University, Guangzhou 510642, China

Hao Wang – National Engineering Research Center for Breeding Swine Industry, Department of Animal Genetics, Breeding and Reproduction, College of Animal Science, Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, State Key Laboratory Swine and Poultry Breeding Industry, and National and local joint Engineering Research Center for Livestock and Poultry Breeding Industry, South China Agricultural University, Guangzhou 510642, China

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jafc.3c09647>

Author Contributions

[†]W.W. and S.W. contributed equally to this work.

Author Contributions

W.W. performed conceptualization, data curation, formal analysis, investigation, writing—original draft, and writing—review and editing. S.W. performed investigation, data curation, and validation. H.W. performed visualization and validation. E.Z. carried out funding acquisition, project administration, and supervision. Z.W. conducted funding acquisition, resource gathering, and writing—review and editing. Z.L. performed conceptualization, funding acquisition, supervision, and writing—review and editing.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by a grant from the Department of Agriculture and Rural Affairs of Guangdong Province, China (grant number: 2023KJ126), and a grant from the Department of Science and Technology of Guangdong Province, China (grant number: 2018B030313011).

ABBREVIATIONS

BTF3 RNA polymerase B transcription factor 3
DAPs differentially abundant proteins
DDA data-dependent acquisition
DIA data-independent acquisition
DTT dithiothreitol
EGF epidermal growth factors
ECM extracellular matrix
EDTA ethylenediaminetetraacetic acid
FW 1d day 1 of forced weaning
FW 3d day 3 of forced weaning

FDR false discovery rate
GO Gene Ontology
HCD higher-energy collisional dissociation
L12d day 12 of lactation
MECs mammary epithelial cells
MIT maximal injection time
MCM3 minichromosome maintenance protein 3
NCE normalized collision energy
PPI protein–protein interaction
PMSF phenylmethylsulfonyl fluoride
PCA principle component analysis
P16d day 16 of pregnancy
SPF specific-pathogen-free
SEM standard error of the mean
V mature virgin

REFERENCES

- (1) Inman, J. L.; Robertson, C.; Mott, J. D.; Bissell, M. J. Mammary gland development: cell fate specification, stem cells and the microenvironment. *Development* **2015**, *142* (6), 1028–1042.
- (2) Peaker, M. The mammary gland in mammalian evolution: a brief commentary on some of the concepts. *J. Mammary Gland Biol. Neoplasia* **2002**, *7* (3), 347–353.
- (3) Hannan, F. M.; Elajna, T.; Vandenberg, L. N.; Kennedy, S. H.; Thakker, R. V. Hormonal regulation of mammary gland development and lactation. *Nat. Rev. Endocrinol* **2023**, *19* (1), 46–61.
- (4) Sankar, M. J.; Sinha, B.; Chowdhury, R.; Bhandari, N.; Taneja, S.; Martinez, J.; Bahl, R. Optimal breastfeeding practices and infant and child mortality: a systematic review and meta-analysis. *Acta Paediatr* **2015**, *104* (467), 3–13.
- (5) Victora, C. G.; Bahl, R.; Barros, A. J.; França, G. V.; Horton, S.; Krasevec, J.; Murch, S.; Sankar, M. J.; Walker, N.; Rollins, N. C. Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *Lancet* **2016**, *387* (10017), 475–490.
- (6) Gunderson, E. P.; Lewis, C. E.; Lin, Y.; Sorel, M.; Gross, M.; Sidney, S.; Jacobs, D. R., Jr; Shikany, J. M.; Quesenberry, C. P., Jr Lactation Duration and Progression to Diabetes in Women Across the Childbearing Years: The 30-Year CARDIA Study. *JAMA Intern Med* **2018**, *178* (3), 328–337.
- (7) Kotsopoulos, J.; Gronwald, J.; McCuaig, J. M.; Karlan, B. Y.; Eisen, A.; Tung, N.; Bordeleau, L.; Senter, L.; Eng, C.; Couch, F.; et al. Breastfeeding and the risk of epithelial ovarian cancer among women with a BRCA1 or BRCA2 mutation. *Gynecol Oncol* **2020**, *159* (3), 820–826.
- (8) Xuan, R.; Wang, J.; Zhao, X.; Li, Q.; Wang, Y.; Du, S.; Duan, Q.; Guo, Y.; Ji, Z.; Chao, T. Transcriptome Analysis of Goat Mammary Gland Tissue Reveals the Adaptive Strategies and Molecular Mechanisms of Lactation and Involution. *Int. J. Mol. Sci.* **2022**, *23* (22), 14424.
- (9) Bach, K.; Pensa, S.; Grzelak, M.; Hadfield, J.; Adams, D. J.; Marioni, J. C.; Khaled, W. T. Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. *Nat. Commun.* **2017**, *8* (1), 2128.
- (10) Twigger, A.-J.; Engelbrecht, L. K.; Bach, K.; Schultz-Pernice, I.; Pensa, S.; Stenning, J.; Petricca, S.; Scheel, C. H.; Khaled, W. T. Transcriptional changes in the mammary gland during lactation revealed by single cell sequencing of cells from human milk. *Nat. Commun.* **2022**, *13* (1), 562.
- (11) Davies, C. R.; Morris, J. S.; Griffiths, M. R.; Page, M. J.; Pitt, A.; Stein, T.; Gusterson, B. A. Proteomic analysis of the mouse mammary gland is a powerful tool to identify novel proteins that are differentially expressed during mammary development. *Proteomics* **2006**, *6* (21), 5694–5704.
- (12) Honvo-Houéto, E.; Truchet, S. Indirect Immunofluorescence on Frozen Sections of Mouse Mammary Gland. *J. Visualized Exp.* **2015**, *106*, 53179.

- (13) Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **2011**, *10* (4), 1794–1805.
- (14) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367–1372.
- (15) Wu, T.; Hu, E.; Xu, S.; Chen, M.; Guo, P.; Dai, Z.; Feng, T.; Zhou, L.; Tang, W.; Zhan, L.; et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2021**, *2* (3), No. 100141.
- (16) Kumar, L.; Futschik, M. E. Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics* **2007**, *2* (1), 5–7.
- (17) von Mering, C.; Jensen, L. J.; Snel, B.; Hooper, S. D.; Krupp, M.; Foglierini, M.; Jouffre, N.; Huynen, M. A.; Bork, P. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res.* **2005**, *33*, D433–D437.
- (18) Sakamoto, K.; Wehde, B. L.; Yoo, K. H.; Kim, T.; Rajbhandari, N.; Shin, H. Y.; Triplett, A. A.; Rädler, P. D.; Schuler, F.; Villunger, A.; et al. Janus Kinase 1 Is Essential for Inflammatory Cytokine Signaling and Mammary Gland Remodeling. *Mol. Cell. Biol.* **2016**, *36* (11), 1673–1690.
- (19) Bionaz, M.; Loor, J. J. Gene networks driving bovine mammary protein synthesis during the lactation cycle. *Bioinf. Biol. Insights* **2011**, *5*, 83–98.
- (20) Mapes, J.; Li, Q.; Kannan, A.; Anandan, L.; Laws, M.; Lydon, J. P.; Bagchi, I. C.; Bagchi, M. K. CUZD1 is a critical mediator of the JAK/STAT5 signaling pathway that controls mammary gland development during pregnancy. *PLoS Genet* **2017**, *13* (3), No. e1006654.
- (21) Asch, H. L.; Asch, B. B. Expression of keratins and other cytoskeletal proteins in mouse mammary epithelium during the normal developmental cycle and primary culture. *Dev. Biol.* **1985**, *107* (2), 470–482.
- (22) Lin, Z.; Zhang, C.; Zhang, M.; Xu, D.; Fang, Y.; Zhou, Z.; Chen, X.; Qin, N.; Zhang, X. Targeting cadherin-17 inactivates Ras/Raf/MEK/ERK signaling and inhibits cell proliferation in gastric cancer. *PLoS One* **2014**, *9* (1), No. e85296.
- (23) Zhao, Y.; Liu, H.; Fan, M.; Miao, Y.; Zhao, X.; Wei, Q.; Ma, B. G protein-coupled receptor 30 mediates cell proliferation of goat mammary epithelial cells via MEK/ERK&PI3K/AKT signaling pathway. *Cell Cycle* **2022**, *21* (19), 2027–2037.
- (24) Fata, J. E.; Chaudhary, V.; Khokha, R. Cellular turnover in the mammary gland is correlated with systemic levels of progesterone and not 17 β -estradiol during the estrous cycle. *Biol. Reprod.* **2001**, *65* (3), 680–688.
- (25) Wang, Y.; Wu, J.; Xia, S. W.; Zhao, F.; Ding, Q.; Ye, X. M.; Zhong, J. F.; Chen, K. L.; Wang, H. L. miR-27a-3p relieves heat stress-induced mitochondrial damage and aberrant milk protein synthesis through MEK/ERK pathway in BMECs. *Cell Stress Chaperones* **2023**, *28* (3), 265–274.
- (26) Oakes, S. R.; Hilton, H. N.; Ormandy, C. J. The alveolar switch: coordinating the proliferative cues and cell fate decisions that drive the formation of lobuloalveoli from ductal epithelium. *Breast Cancer Res.* **2006**, *8* (2), 207.
- (27) Richert, M. M.; Schwertfeger, K. L.; Ryder, J. W.; Anderson, S. M. An atlas of mouse mammary gland development. *J. Mammary Gland Biol. Neoplasia* **2000**, *5* (2), 227–241.
- (28) Anderson, S. M.; Rudolph, M. C.; McManaman, J. L.; Neville, M. C. Key stages in mammary gland development. Secretory activation in the mammary gland: it's not just about milk protein synthesis! *Breast Cancer Res.* **2007**, *9* (1), 204.
- (29) Lund, L. R.; Rømer, J.; Thomasset, N.; Solberg, H.; Pyke, C.; Bissell, M. J.; Danø, K.; Werb, Z. Two distinct phases of apoptosis in mammary gland involution: proteinase-independent and -dependent pathways. *Development* **1996**, *122* (1), 181–193.
- (30) Hennigar, S. R.; Seo, Y. A.; Sharma, S.; Soybel, D. I.; Kelleher, S. L. ZnT2 is a critical mediator of lysosomal-mediated cell death during early mammary gland involution. *Sci. Rep.* **2015**, *5*, 8033.
- (31) McNally, S.; Stein, T. Overview of Mammary Gland Development: A Comparison of Mouse and Human. *Methods Mol. Biol.* **2017**, *1501*, 1–17.
- (32) Haricharan, S.; Li, Y. STAT signaling in mammary gland differentiation, cell survival and tumorigenesis. *Mol. Cell. Endocrinol.* **2014**, *382* (1), 560–569.
- (33) Humphreys, R. C.; Brier, B.; Zhao, L.; Raz, R.; Levy, D.; Hennighausen, L. Deletion of Stat3 blocks mammary gland involution and extends functional competence of the secretory epithelium in the absence of lactogenic stimuli. *Endocrinology* **2002**, *143* (9), 3641–3650.
- (34) Jena, M. K.; Jaswal, S.; Kumar, S.; Mohanty, A. K. Molecular mechanism of mammary gland involution: An update. *Dev. Biol.* **2019**, *445* (2), 145–155.
- (35) Baxter, F. O.; Neoh, K.; Tevendale, M. C. The beginning of the end: death signaling in early involution. *J. Mammary Gland Biol. Neoplasia* **2007**, *12* (1), 3–13.
- (36) Kritikou, E. A.; Sharkey, A.; Abell, K.; Came, P. J.; Anderson, E.; Clarkson, R. W.; Watson, C. J. A dual, non-redundant, role for LIF as a regulator of development and STAT3-mediated cell death in mammary gland. *Development* **2003**, *130* (15), 3459–3468.
- (37) Nguyen, A. V.; Pollard, J. W. Transforming growth factor beta3 induces cell death during the first stage of mammary gland involution. *Development* **2000**, *127* (14), 3107–3118.
- (38) Macias, H.; Hinck, L. Mammary gland development. *Wiley Interdiscip. Rev. Dev. Biol.* **2012**, *1* (4), 533–557.
- (39) Biswas, S. K.; Banerjee, S.; Baker, G. W.; Kuo, C. Y.; Chowdhury, I. The Mammary Gland: Basic Structure and Molecular Signaling during Development. *Int. J. Mol. Sci.* **2022**, *23* (7), 3883.
- (40) Stoerber, K.; Tlsty, T. D.; Happerfield, L.; Thomas, G. A.; Romanov, S.; Bobrow, L.; Williams, E. D.; Williams, G. H. DNA replication licensing and human cell proliferation. *J. Cell Sci.* **2001**, *114* (Pt 11), 2027–2041.
- (41) Ewida, M. A.; Ewida, H. A.; Ahmed, M. S.; Allam, H. A.; ElBagary, R. I.; George, R. F.; Georgey, H. H.; El-Subbagh, H. I. 3-Methyl-imidazo[2,1-b]thiazole derivatives as a new class of antifolates: Synthesis, in vitro/in vivo bio-evaluation and molecular modeling simulations. *Bioorg. Chem.* **2021**, *115*, No. 105205.
- (42) Topper, Y. J.; Freeman, C. S. Multiple hormone interactions in the developmental biology of the mammary gland. *Physiol. Rev.* **1980**, *60* (4), 1049–1106.
- (43) Yamaji, D.; Kang, K.; Robinson, G. W.; Hennighausen, L. Sequential activation of genetic programs in mouse mammary epithelium during pregnancy depends on STAT5A/B concentration. *Nucleic Acids Res.* **2013**, *41* (3), 1622–1636.
- (44) Wilde, C. J.; Knight, C. H.; Flint, D. J. Control of milk secretion and apoptosis during mammary involution. *J. Mammary Gland Biol. Neoplasia* **1999**, *4* (2), 129–136.
- (45) Watson, C. J. Involution: apoptosis and tissue remodelling that convert the mammary gland from milk factory to a quiescent organ. *Breast Cancer Res.* **2006**, *8* (2), 203.
- (46) Valsalan, J.; Sadan, T.; Venkatachalapathy, T.; Anilkumar, K.; Aravindakshan, T. V. Identification of novel single-nucleotide polymorphism at exon1 and 2 region of B4GALT1 gene and its association with milk production traits in crossbred cattle of Kerala. *India. Anim. Biotechnol.* **2022**, *33* (6), 1056–1064.
- (47) Song, Y.; Zhang, J.; Jiang, C.; Song, X.; Wu, H.; Zhang, J.; Raza, S. H. A.; Zhang, L.; Zhang, L.; Cai, B.; et al. FOXO1 regulates the formation of bovine fat by targeting CD36 and STEAP4. *Int. J. Biol. Macromol.* **2023**, *248*, No. 126025.
- (48) Hovey, R. C.; Aimo, L. Diverse and active roles for adipocytes during mammary gland growth and function. *J. Mammary Gland Biol. Neoplasia* **2010**, *15* (3), 279–290.
- (49) Rudolph, M. C.; Neville, M. C.; Anderson, S. M. Lipid synthesis in lactation: diet and the fatty acid switch. *J. Mammary Gland Biol. Neoplasia* **2007**, *12* (4), 269–281.

- (50) Zwick, R. K.; Rudolph, M. C.; Shook, B. A.; Holtrup, B.; Roth, E.; Lei, V.; Van Keymeulen, A.; Seewaldt, V.; Kwei, S.; Wysolmerski, J.; et al. Adipocyte hypertrophy and lipid dynamics underlie mammary gland remodeling after lactation. *Nat. Commun.* **2018**, *9* (1), 3592.
- (51) Watson, C. J.; Kreuzaler, P. A. Remodeling mechanisms of the mammary gland during involution. *Int. J. Dev. Biol.* **2011**, *55* (7–9), 757–762.
- (52) Kobayashi, K.; Oyama, S.; Kuki, C.; Tsugami, Y.; Matsunaga, K.; Suzuki, T.; Nishimura, T. Distinct roles of prolactin, epidermal growth factor, and glucocorticoids in β -casein secretion pathway in lactating mammary epithelial cells. *Mol. Cell. Endocrinol.* **2017**, *440*, 16–24.
- (53) Aber, R.; Chan, W.; Mugisha, S.; Jerome-Majewska, L. A. Transmembrane emp24 domain proteins in development and disease. *Genet. Res.* **2019**, *101*, e14.
- (54) Park, H.; Seo, S. K.; Sim, J. R.; Hwang, S. J.; Kim, Y. J.; Shin, D. H.; Jang, D. G.; Noh, S. H.; Park, P. G.; Ko, S. H.; et al. TMED3 Complex Mediates ER Stress-Associated Secretion of CFTR, Pendrin, and SARS-CoV-2 Spike. *Adv. Sci. (Weinh)* **2022**, *9* (24), e2105320.
- (55) Ma, Y. F.; Batistel, F.; Xu, T. L.; Han, L. Q.; Bucktrout, R.; Liang, Y.; Coleman, D. N.; Parys, C.; Loor, J. J. Phosphorylation of AKT serine/threonine kinase and abundance of milk protein synthesis gene networks in mammary tissue in response to supply of methionine in periparturient Holstein cows. *J. Dairy Sci.* **2019**, *102* (5), 4264–4274.
- (56) Shimobayashi, M.; Hall, M. N. Making new contacts: the mTOR network in metabolism and signalling crosstalk. *Nat. Rev. Mol. Cell Biol.* **2014**, *15* (3), 155–162.
- (57) Chat, S.; Layani, S.; Mahaut, C.; Henry, C.; Chanat, E.; Truchet, S. Characterisation of the potential SNARE proteins relevant to milk product release by mouse mammary epithelial cells. *Eur. J. Cell Biol.* **2011**, *90* (5), 401–413.
- (58) Rios, A. C.; Fu, N. Y.; Jamieson, P. R.; Pal, B.; Whitehead, L.; Nicholas, K. R.; Lindeman, G. J.; Visvader, J. E. Essential role for a novel population of binucleated mammary epithelial cells in lactation. *Nat. Commun.* **2016**, *7*, 11400.
- (59) Naylor, M. J.; Oakes, S. R.; Gardiner-Garden, M.; Harris, J.; Blazek, K.; Ho, T. W.; Li, F. C.; Wynick, D.; Walker, A. M.; Ormandy, C. J. Transcriptional changes underlying the secretory activation phase of mammary gland development. *Mol. Endocrinol.* **2005**, *19* (7), 1868–1883.
- (60) Qian, X.; Zhao, F. Q. Current major advances in the regulation of milk protein gene expression. *Crit. Rev. Eukaryot. Gene Expr.* **2014**, *24* (4), 357–378.
- (61) Neville, M. C.; McFadden, T. B.; Forsyth, I. Hormonal regulation of mammary differentiation and milk secretion. *J. Mammary Gland Biol. Neoplasia* **2002**, *7* (1), 49–66.
- (62) Kawecka-Grochowska, E.; Zalewska, M.; Kapusta, A.; Ząbek, T.; Rzewuska, M.; Petrykowski, S.; Bagnicka, E. Transcripts and protein levels of CSN1S1 and CSN3 genes in dairy cattle mammary gland secretory tissue during chronic staphylococcal infection. *J. Dairy Res.* **2021**, *88* (1), 73–77.
- (63) Fu, N. Y.; Rios, A. C.; Pal, B.; Soetanto, R.; Lun, A. T.; Liu, K.; Beck, T.; Best, S. A.; Vaillant, F.; Bouillet, P.; et al. EGF-mediated induction of Mcl-1 at the switch to lactation is essential for alveolar cell survival. *Nat. Cell Biol.* **2015**, *17* (4), 365–375.
- (64) Kisseleva, T.; Bhattacharya, S.; Braunstein, J.; Schindler, C. W. Signaling through the JAK/STAT pathway, recent advances and future challenges. *Gene* **2002**, *285* (1–2), 1–24.
- (65) Rodig, S. J.; Meraz, M. A.; White, J. M.; Lampe, P. A.; Riley, J. K.; Arthur, C. D.; King, K. L.; Sheehan, K. C.; Yin, L.; Pennica, D.; et al. Disruption of the Jak1 gene demonstrates obligatory and nonredundant roles of the Jaks in cytokine-induced biologic responses. *Cell* **1998**, *93* (3), 373–383.
- (66) Khaled, W. T.; Read, E. K.; Nicholson, S. E.; Baxter, F. O.; Brennan, A. J.; Came, P. J.; Sprigg, N.; McKenzie, A. N.; Watson, C. J. The IL-4/IL-13/Stat6 signalling pathway promotes luminal mammary epithelial cell development. *Development* **2007**, *134* (15), 2739–2750.
- (67) Oliver, C. H.; Khaled, W. T.; Frend, H.; Nichols, J.; Watson, C. J. The Stat6-regulated KRAB domain zinc finger protein Zfp157 regulates the balance of lineages in mammary glands and compensates for loss of Gata-3. *Genes Dev.* **2012**, *26* (10), 1086–1097.
- (68) Ma, J.; Chen, T.; Wu, S.; Yang, C.; Bai, M.; Shu, K.; Li, K.; Zhang, G.; Jin, Z.; He, F.; Hermjakob, H.; Zhu, Y. iProX: an integrated proteome resource. *Nucleic acids res.* **2019**, *47* (D1), D1211–D1217.
- (69) Chen, T.; Ma, J.; Liu, Y.; Chen, Z.; Xiao, N.; Lu, Y.; Fu, Y.; Yang, C.; Li, M.; Wu, S.; et al. iProX in 2021: connecting proteomics data sharing with big data. *Nucleic Acids Res.* **2022**, *50* (D1), D1522–d1527.

RESEARCH

Open Access



Genome-wide detection of multiple variants associated with teat number in French Yorkshire pigs

Danyang Lin^{1†}, Yibin Qiu^{1†}, Fuchen Zhou¹, Xuehua Li¹, Shaoxiong Deng¹, Jisheng Yang¹, Qiaoer Chen¹, Gengyuan Cai^{1,2,3}, Jie Yang^{1,3}, Zhenfang Wu^{1,2,3,4*} and Enqin Zheng^{1*}

Abstract

Background Teat number is a vital reproductive trait in sows, crucial for providing immunity and nutrition to piglets during lactation. However, “missing heritability” in Single Nucleotide Polymorphism (SNP)-based Genome-Wide Association Studies (GWAS) has led to an increasing focus on structural variations in the genetic analysis of complex biological traits.

Results In this study, we generated a comprehensive CNV map in a population of French Yorkshire pigs ($n=644$) and identified 429 CNVRs. Notably, 44% (189 CNVRs) of these were detected for the first time. Subsequently, we conducted GWAS for teat number in the French Yorkshire pig population using both 80K chip and its imputed data, as well as a GWAS analysis based on CNV regions (CNVR). Interestingly, 80K chip GWAS identified two SNPs located on *Sus scrofa* chromosome 5 (SSC5) that were simultaneously associated with Total Teat Number (TTN), Left Teat Number (LTN), and Right Teat Number (RTN). The leading SNP (WU_10.2_5_76130558) explained 3.33%, 2.69%, and 2.67% of the phenotypic variance for TTN, LTN, and RTN, respectively. Moreover, through imputed GWAS, we successfully identified 30 genetic variants associated with TTN located within the 73.22–73.30 Mb region on SSC5. The two SNPs identified in the 80K chip GWAS were also located in this region. In addition, CNVR-based GWAS revealed three significant CNVRs associated with TTN. Finally, through gene annotation, we pinpointed two candidate genes, *TRIM66* and *PRICKLE1*, which are related to diverse processes such as breast cancer and abnormal vertebral development.

Conclusions Our research provides an in-depth analysis of the complex genetic structure underlying teat number, contributing to the genetic enhancement of sows with improved reproductive performance and, ultimately, bolstering the economic benefits of swine production enterprises.

Keywords Teat number, Copy number variation, Genome-wide association study, French Yorkshire pigs

[†]Danyang Lin and Yibin Qiu contributed equally to this work.

*Correspondence:

Zhenfang Wu

wzfemail@163.com

Enqin Zheng

eqzheng@scau.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Teat number, as an important reproduction trait, has a large effect on the lactation ability of the sows, thus teat number may directly affect the weight gain of piglets [1, 2]. However, the complex genetic architecture of teat number makes its molecular mechanism unclear. With the rapid development and application of high-throughput sequencing technology, genome-wide association analysis (GWAS) combined with molecular marker technology is considered a powerful approach for dissecting the genetic architecture of complex traits in livestock [3–5]. Several previous GWAS researches indicate that *VRTN* [6] and *ABCD4* [7] as candidate genes may regulate teat upgrowth, but traditional GWAS based on single nucleotide polymorphisms (SNP) have accounted for only part of the total heritability [8].

Some of the missing heritability has been attributed to copy number variation (CNV) in humans [9]. CNV refers to the structural variations of DNA segments in the range of 50 bp to several Mb compared with a reference genome, and it is widespread distributed in the genome [10]. Overlapping CNVs are combined into big regions known as the copy number variation regions (CNVR) [11]. CNVs may lead to phenotypic variation and disease by altering gene structure, gene regulation, and exposure to recessive deleterious genes [12]. As researchers revealed the existence of large-scale copy number variation in the human genome [13, 14], CNV research was gradually carried out in various fields [15–17]. For example, Chen et al. [18] identified a CNV affecting the *MSRB3* gene that increases pig ear size through the mechanism of Mir-584-5p; Wang et al. [19] found that the gain status at CNVR will decrease total number born and number born alive in large white sows. Overall, research into novel CNVs of pigs can capture part of missing heritability from SNP-based GWAS and explain more genomic structural variations.

In this study, multi-dataset GWAS were conducted for teat number in the French Yorkshire pig population. Our research aimed to identify genetic variants and candidate genes associated with teat number in pigs and to elucidate the potential molecular genetic mechanisms. Additionally, the genome-wide CNV detection provides a valuable complement for CNV map of the French Yorkshire pig genome.

Results

Phenotype and heritability statistics

In this study, we analyzed three traits including Total Teat Number (TTN), Left Teat Number (LTN), and Right Teat Number (RTN) in 644 French Yorkshire pigs (Table 1). The average numbers (mean ± standard deviation) of TTN, LTN, and RTN were 14.10 ± 0.92 , 7.04 ± 0.50 and 7.06 ± 0.49 , respectively. The coefficients of variation (C.V.) of all three traits were over 6.50%. Besides, we estimated the SNP based heritability (h^2) for the three traits based on the 53,869 SNPs, and the genomic h^2 ranged from 0.14 ± 0.06 to 0.17 ± 0.07 , suggesting that teat number belongs to low heritability trait.

Detection of genome-wide CNVs

A total of 8,746 CNVs (583 losses and 8163 gains) were detected using PennCNV software v1.0.5 [20], with an average of 13.48 per individual, ranging from two to 32 (Fig. 1a). Compared with losses, CNV gains occurred more frequently in individuals (Fig. 1a). The length of CNVs pigs ranged from 10.4 kb to 1.25 Mb with an average of 128.80 kb, and the median length of gains was 99.85 kb, which was longer than that of loss (82.63 kb, Fig. 1b). As shown in Table S1, all CNVs were merged into 429 CNVRs, including 103 losses, 296 gains, and 30 mixed events (gains and losses occurring in the same region). The total length of CNVRs was 66.78 Mb, occupying 2.95% of the total length of the pig autosomal genome (*Sus scrofa* version 11.1).

Table 1 The statistics for the phenotypes of teat number

Trait ^a	N ^b	Mean (± SD) ^c	Min ^d	Max ^e	C.V. (%) ^f	h^2 (SE) ^g
TTN	629	14.10 ± 0.92	10	18	6.52	0.17 ± 0.07
LTN	631	7.04 ± 0.50	5	9	7.14	0.14 ± 0.06
RTN	630	7.06 ± 0.49	5	9	6.90	0.17 ± 0.07

^a TTN Total teat number, LTN Left teat number, RTN Right teat number

^b Number of animals (N)

^c Mean (standard deviation)

^d Minimum (min)

^e Maximum (max)

^f Coefficient of variation (C.V.)

^g Heritability (standard error)

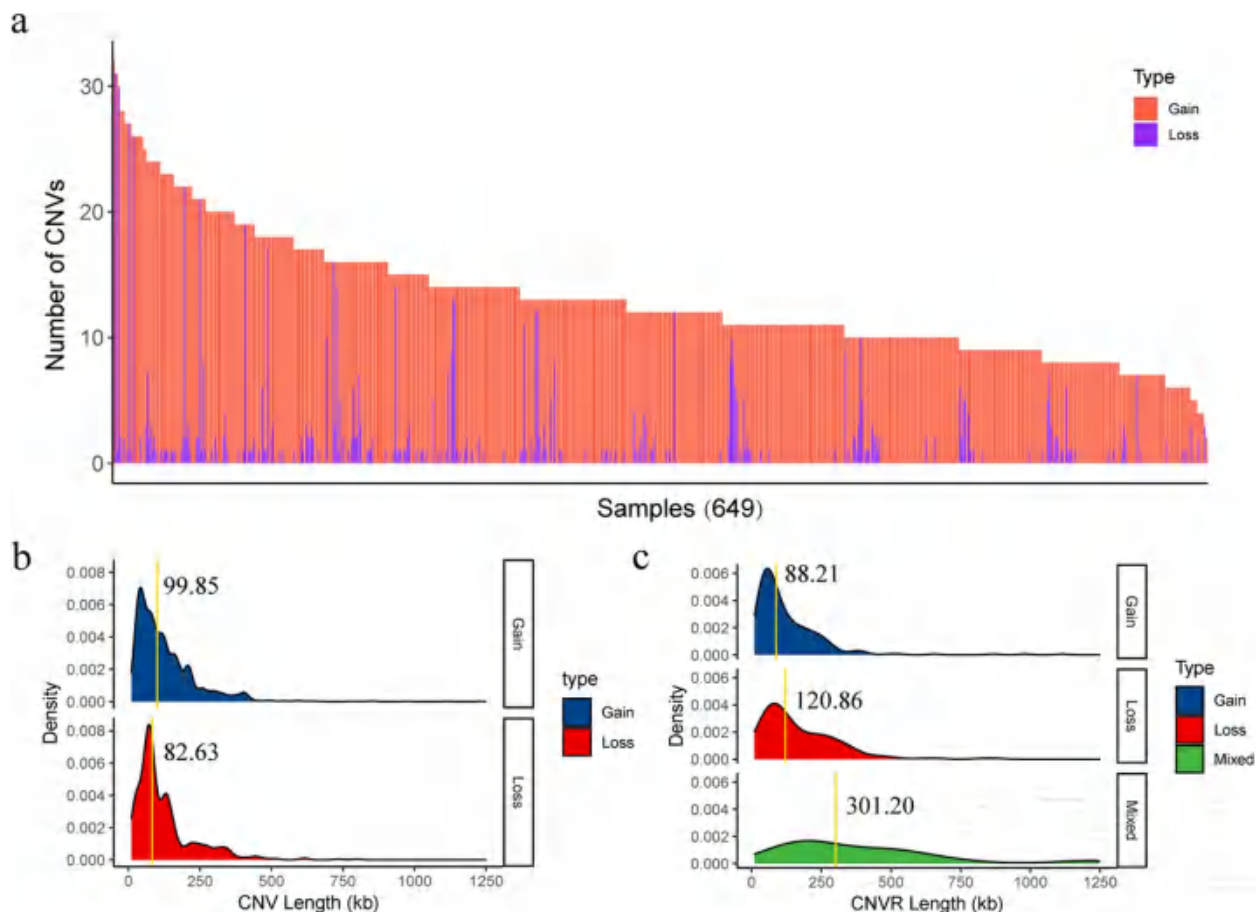


Fig. 1 Number and length of different CNV and CNVR type. **a** Number of CNVs of each type in each individual. **b** Length distribution for each CNV type. The gold line indicates the median length of each CNV type. **c** Length distribution for each CNVR type. The gold line indicates the median length of each CNVR type

The length of CNVRs ranged from 10.40 kb to 1.25 Mb with an average of 155.66 kb and the median length of gains (88.21 kb) was shorter than that of losses (120.86 kb) and of mixed events (301.20 kb, Fig. 1c). In addition, we observed that the majority of CNVs and CNVRs were under 500 kb in size.

Figure 2a and Table 2 illustrate the number and proportion of CNVRs distributed on autosomal. The number of CNVRs ranged from 11 in chromosome 18 (SSC18) to 36 in SSC2, accounting for 2.56% and 8.39% of the total proportion of CNVRs, respectively. In addition, CNVRs on SSC12 were the most dense, with an average distance between adjacent CNVRs of 2.12 Mb. We also found that the density distribution of the SNP and CNVR were remarkably consistent, suggesting that increasing SNP density contributes to the detection rate of CNVR [20] (Fig. 2a).

Gene content of CNVRs

A total of 1,558 genes from the Ensembl annotation of the *Sscrofa* 11.1 genome were identified to be overlapping with our detected 429 CNVRs, including 870 known genes and 688 unknown genes. Of these, 74.33% were protein-coding genes, 16.11% were long noncoding RNA (lncRNA), and others belonged to pseudogenes, small nuclear RNA (snRNA), microRNA (miRNA), small nucleolar RNA (snoRNA), processed pseudogenes, miscellaneous RNA (miscRNA), small cajal body-specific RNA (scaRNA), and T cell receptors (TR) V gene.

To further study the 1558 genes contended in CNVRs, we performed GO analysis and KEGG pathway analysis, as shown in Table S2. Accordingly, the GO analysis showed that genes of the terms of proteolysis, calcium ion binding, and endoplasmic reticulum are dominantly represented in the CNVRs (Fig. S1a), and the KEGG

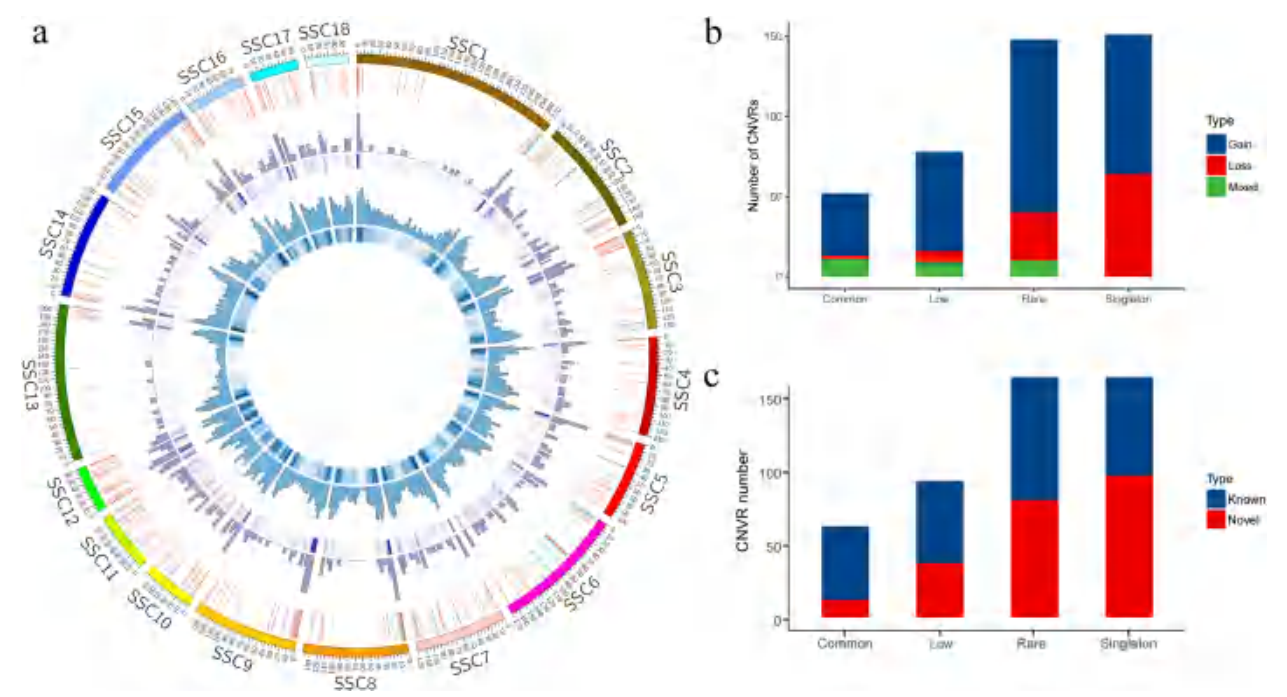


Fig. 2 Comprehensive representation of CNVR maps, frequencies, and known/novel CNVR counts for French Yorkshire in the 18 autosomes. **a** Outer to inner circles: chromosome name; genomic location (in Mb); arrangement of CNVRs on the genome (gain in red, loss in green, mixed in blue); density histogram of CNVRs in 5 Mb bin (purple); density heatmap of CNVRs in 5 Mb bin; density histogram of SNPs in 5 Mb bin (blue); and density heatmap of SNPs in 5 Mb bin (blue). **b** Frequencies of CNVRs in the French Yorkshire population. **c** Known and novel CNVR counts categorized by frequencies

Table 2 Chromosome distribution of all 429 CNVRs in the pig autosomes

Chr	Chr length (kb)	CNVR counts	Length of CNVR (kb)	Coverage (%)	Max size (kb)	Average size (kb)	Min size (kb)
1	274,330.53	26	4,527.88	1.65	526.39	155.71	13.09
2	151,935.99	36	4,621.77	3.04	553.78	87.92	19.52
3	132,848.91	29	4,486.68	3.38	1114.33	99.61	21.2
4	130,910.91	22	2,674.49	2.04	375.83	97.04	16.65
5	104,526.01	23	4,848.07	4.64	643.54	196.31	21.53
6	170,843.59	32	7,883.07	4.61	1,248.48	117.01	15.53
7	121,844.10	25	3,897.32	3.2	447.9	125.94	33.73
8	138,966.24	22	3,807.39	2.74	1,170.07	98.87	14.66
9	139,512.08	27	3,163.48	2.27	380.96	81.85	19.07
10	69,359.45	19	1,605.59	2.31	214.26	80.85	10.4
11	79,169.98	22	3,739.16	4.72	581.58	115.83	30.47
12	61,602.75	30	4,667.07	7.58	869.96	102.23	31.75
13	208,334.59	17	2,845.53	1.37	763.66	141.64	20.47
14	141,755.45	21	3,118.45	2.2	495.34	101.22	24.92
15	140,412.72	32	4,454.12	3.17	656.86	86.06	15.9
16	79,944.28	17	2,819.25	3.53	559.87	135.21	15.54
17	63,494.08	18	2,563.14	4.04	528.22	95.6	15.49
18	55,982.97	11	1,057.43	1.89	236.16	84.44	35.64

pathway analysis revealed that these genes are mainly represented in the pathway of endocytosis and estrogen signaling pathway (Fig. S1b). Compared with the reported quantitative trait loci (QTLs) in pigs, a total of 419 (97.67%) CNVRs included in or partially overlapping with 17674 QTLs (Table S3), which are associated with a variety of traits, such as teat number, average daily gain, and body weight. Among these QTLs, a number of 408 QTLs were associated with teat number.

Comparison of CNVRs detected in previous studies

The CNVR data set detected in this study was compared with the data of previous studies on CNVR [19, 21–32], as shown in Table S4. The results of Zheng et al. [31] have the highest overlap rate, with 609 CNVRs, while those of Wang et al. [22] have the lowest overlap rate, with only three CNVRs, indicating that there are still a large number of CNVs in the pig genome haven't yet been discovered. Further, 189 CNVRs were newly identified, meanwhile, we detected 240 CNVRs overlapping with the previously reported CNVRs.

CNVR frequency in French Yorkshire population

The frequencies of CNVRs in the French Yorkshire population were also calculated and grouped into four categories: singleton (present in one individual), rare (present in more than one individual but with a frequency ≤ 0.01), low ($0.01 < \text{frequency} \leq 0.05$), and common (frequency > 0.05), as shown in Fig. 2b. The singleton (151) accounted for 35.2% of all CNVRs, while there were 148 rare (34.5%), 78 low (18.2%) and 52 common (12.1%) CNVRs. As a result of the 189 novel CNVRs, 42.9%, 35.4%, 16.4% and 5.3% of them belong to singleton, rare, low and common, respectively. We found that the proportion of novel CNVRs was larger when the frequency of identified CNVRs were lower (Fig. 2c). The frequencies of CNVRs ranged from 0.15% (detected in one pig) to 71.76% (detected in 465 pigs) and was concentrated in singleton and rare frequency categories, indicating that CNVRs occur mostly in a few individuals and is hard to measure reliably [33]. Thus, we used 130 CNVRs with a frequency greater than 1% for later GWAS.

SNP-based GWAS results

PCA based on SNPs showed that there was no population stratification in this population (Fig. S2). Additionally, quantile–quantile (Q–Q) plots were used to illustrate the level of potential *P* value inflation (Fig. S3). The genomic inflation factors (λ) of GWAS ranged from 0.976 to 0.992, indicating that there is no obvious evidence of population stratification. Significant SNPs detected through 80K chips GWAS are presented in Fig. 3 and Table 3. We found that two SNPs (WU_10.2_5_76130558 and

WU_10.2_5_76207514) located in SSC5 were simultaneously associated with all of the TTN, LTN, and RTN. The leading SNP (WU_10.2_5_76130558) explained 3.33%, 2.69% and 2.67% phenotypic variance for TTN, LTN and RTN, respectively. Subsequently, we performed a haplotype block analysis and found that these two SNPs are in complete LD and located within a 46 kb haplotype block (5: 73.19 Mb–73.24 Mb), which suggests that mutations near the potential QTL may have essential effect on teat number (Fig. S4).

Furthermore, GWAS conducted by imputed data revealed 30 significant variants for TTN (Fig. S5 and Table S5). The leading variant is 5_73264327_C located on SSC5 ($P=4.54E-07$). Additionally, the 30 variants are located between 73.22 Mb and 73.30 Mb on SSC5, suggesting the presence of potentially significant variants affecting TTN. Notably, both the WU_10.2_5_76130558 and WU_10.2_5_76207514, identified by the 80K chip GWAS, are also within this region.

CNVR-based GWAS results

PCA based on CNVR also shows that there was no population stratification (Fig. S6). To further dissect the genetic basis of the teat number, CNVR-based GWAS were performed on 644 pigs with phenotypic records for TTN, LTN and RTN, respectively. Figure 4 shows the Manhattan plots for TTN, LTN, and RTN obtained from the association analysis, respectively. As shown in Table 4, we identified two CNVRs located on SSC1 and SSC15 that are associated with both TTN and RTN, demonstrating significant multi-effect associations. We also identified one CNVRs associated with LTN, located on SSC9.

Functional analysis of candidate genes

A total of 12 candidate protein-coding genes overlapped with significant CNVRs and located within a 1 Mb region surrounding the significant variants were detected based on the *Sus scrofa* 11.1. Subsequently, we employed the GeneCards, Mouse Genome Informatics databases, and conducted an extensive literature review to explore the functional roles of the identified genes. As a result, we identified *TRIM66* and *PRICKLE1* genes that exhibiting promising associations with teat number based on their known functions and previous research findings.

Discussion

Comparison of CNVR detected in this study with previous studies

In this study, a total of 429 CNVRs were detected using GeneSeek Porcine 80K SNP chip on 649 French Yorkshire pigs, which provides a supreme valuable supplement for the CNV map. The results showed that the gain

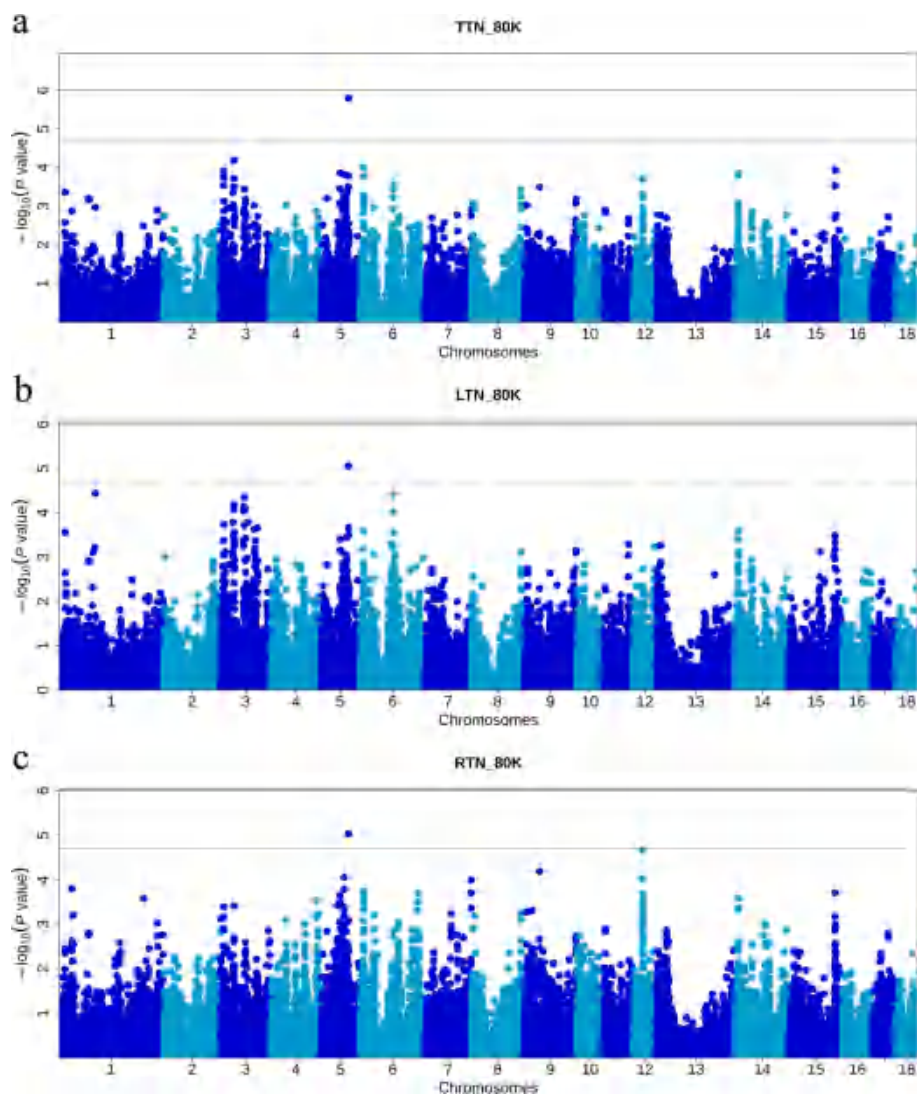


Fig. 3 Manhattan plots of 80K chip GWAS in this population. Manhattan plots consisted of TTN (a), LTN (b) and RTN (c), respectively. The x-axis represents the chromosomes, and the y-axis represents the $-\log_{10}(P\text{-value})$. The solid and dashed lines indicate the 5% genome-wide ($P=9.94\text{E-}07$) and suggestive ($P=1.99\text{E-}05$) Bonferroni-corrected thresholds, respectively

Table 3 Significant SNPs associated with teat number in this population

Trait ^a	Chr	SNP ID	Position (bp) ^b	MAF	P-Value	EPV ^c	Candidate Gene
TTN	5	WU_10.2_5_76130558	73,194,895	0.119	1.60E-06	3.33%	YAF2, ZCRB1, PDZRN4, PPHLN1, PRICKLE1, GXYLT1
		WU_10.2_5_76207514	73,241,384	0.119	1.60E-06		
LTN	5	WU_10.2_5_76130558	73,194,895	0.119	8.69E-06	2.69%	YAF2, ZCRB1, PDZRN4, PPHLN1, PRICKLE1, GXYLT1
		WU_10.2_5_76207514	73,241,384	0.119	8.69E-06		
RTN	5	WU_10.2_5_76130558	73,194,895	0.119	9.27E-06	2.67%	YAF2, ZCRB1, PDZRN4, PPHLN1, PRICKLE1, GXYLT1
		WU_10.2_5_76207514	73,241,384	0.119	9.27E-06		

^a TTN Total teat number, LTN Left teat number, RTN Right teat number
^b SNP position in Ensembl (Position)
^c Explained phenotypic variance

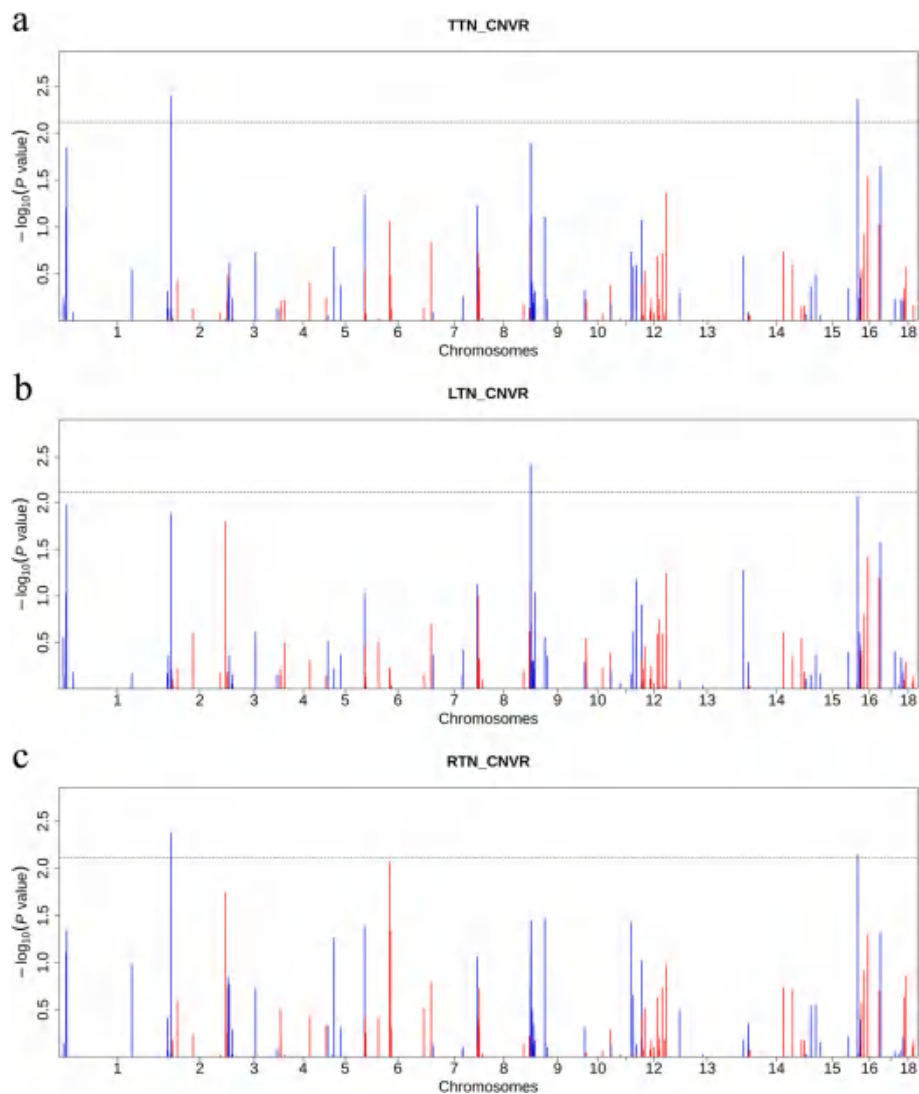


Fig. 4 Manhattan plots of CNVR-based GWAS in this population. Manhattan plots consisted of TTN (a), LTN (b) and RTN (c), respectively. The x-axis represents the chromosomes, and the y-axis represents the $-\log_{10}(P\text{-value})$. The solid and dashed lines indicate the 5% genome-wide ($P=3.85E-04$) and suggestive ($P=7.69E-03$) Bonferroni-corrected thresholds, respectively

Table 4 Significant CNVRs associated with teat number in French Yorkshire pigs

Trait ^a	CNVR_ID	Type	Chr	Start (bp)	End (bp)	Frequency	P-Value	Candidate Gene
TTN	CNVR_23	GAIN	1	271,514,775	271,598,983	3.70%	4.08E-03	RAPGEF1
	CNVR_376	GAIN	15	133,966,827	134,022,507	1.08%	4.30E-03	
LTN	CNVR_217	GAIN	9	654,880	768,919	1.39%	3.88E-03	DENND2B, TRIM66
RTN	CNVR_23	GAIN	1	271,514,775	271,598,983	3.70%	4.17E-03	RAPGEF1
	CNVR_376	GAIN	15	133,966,827	134,022,507	1.08%	7.14E-03	

^a TTN Total teat number, LTN Left teat number, RTN Right teat number

CNVRs was higher than the loss CNVRs, which may be related to the stronger resistance of the genome to duplication than to deletion [34]. In addition, as our previous study, we also found that CNVRs occur more frequently in telomeres (seven of the top 10 largest CNVRs) [35], which ensure the stability and integrity of the genome

and are associated with the replication of genetic material [36].

In previous studies, Wang et al. [27] performed CNVR detection in 12 pigs including nine pig breeds using 1M aCGH and obtained 758 CNVRs (*Sus Scrofa* 10.2), 20 of which are overlapped with our results. Xie et al. [26] detected 172 CNVRs (*Sus Scrofa* 10.2) in 125 pigs using the pig 60K SNPs chip, with only six overlapping with ours. As such, these differences may pertain to various platforms, detection software, algorithms, variety, and quantity of samples, etc. [37–40].

We also compared the sizes of CNVRs detected in different studies. The average length of CNVRs identified in current study is 128.80 kb. In contrast, previous studies using pig SNP chips reported CNVR sizes ranging from 148.99 kb to 1835.44 kb, while those using next-generation sequencing data reported sizes from 4.16 kb to 7.04 kb. The uneven distribution of SNP in Illumina high-density SNP genotyping arrays results in some small CNVs being easily missed in detection [20, 41]. So we concluded that increasing the marker density could improve the detection efficiency and accuracy of CNV, especially small fragment CNV.

Candidate genes associated with teat number

In this study, the SNP-based GWAS and CNVR-based GWAS did not detect overlapping signals. SNPs typically affect single nucleotide changes, whereas CNVs involve large segments of DNA with duplications or deletions that can encompass multiple genes or non-coding regions. Therefore, SNPs and CNVRs may regulate phenotypes by influencing different biological processes or pathways.

Despite this, through multi-dataset GWASs, we identified two candidate genes associated with teat number, namely *TRIM66* and *PRICKLE1*. These two genes have known associations with various processes that could potentially influence teat number, such as breast cancer and abnormal vertebral development.

TRIM66 (Tripartite Motif Containing 66) is a protein-coding gene and a member of the tri-motif protein family. Ning et al. [42] integrated several datasets and software to perform a comprehensive analysis of the expression pattern in *TRIMs* and found that *TRIM66* is significantly downregulated in breast cancer. Zhang et al. [43] also discovered that knocking down *TRIM66* inhibits the proliferation of breast cancer cells. Similar to our findings, other GWAS studies on teat number have also identified several candidate genes with functions related to breast cancer [5, 44]. From a biological perspective, mammary gland development is fundamental to teat formation. Genes that affect breast cancer may also influence normal mammary gland

development and function, thereby indirectly affecting teat formation. Although our current results require extensive validation, they hold promise for enhancing our understanding of the underlying regulatory mechanisms of teat number in pigs.

PRICKLE1, located on SSC5 at 73.71–73.83 Mb, approximately 416.48 kb from the variant region related to teat number in our results. *PRICKLE1* encodes a nuclear receptor and is associated with abnormal vertebral development [45] and other processes. Previous research suggests that *PRICKLE1* is involved in the Wnt signaling pathway [45], which is crucial for mammary gland and thymus development [46]. In previous studies, many strong candidate genes associated with teat number have also been demonstrated to be related to vertebral development, such as *VRTN* and *MKX* [6, 47]. Additionally, previous researches indicated that *VRTN* may be an important gene affecting teat number [48–50], but Zhuang et al. [7] showed that *VRTN* is not the most significant gene affecting teat number, and genetic heterogeneity of its insertion may exist in different populations. However, the signal of *VRTN* was not scanned in our analysis results, which may be due to the differences in varieties, the small size of the population and the high degree of inbreeding, leading to the low minor allele frequency value and finally was eliminated in the quality control.

Heritability and asymmetry of teat number in pigs

In this study, the heritability estimates for TTN, LTN, and RTN were 0.17 ± 0.07 , 0.14 ± 0.06 , and 0.17 ± 0.07 , respectively, with LTN having a lower heritability than TTN and RTN. Similar results were observed in previous studies. Wei et al. [51] investigated teat number traits in a large sample of pigs, defining not only TTN, LTN, and RTN, but also the maximum per side of teat number (TNMPS), teat number symmetry (TNSYM), and the difference between sides of teat number (TNUMD). They found that the heritabilities for TTN and RTN were moderate (0.142 – 0.146), whereas those for LTN, TNMPS, TNSYM, and TNUMD were lower (0.048 – 0.097). After accounting for epistatic effects, the heritability for RTN decreased (0.047), and those for LTN and TNMPS were moderate (0.107 – 0.126). Additionally, studies by Li et al. [5] and Liu et al. [44] also found differing heritability estimates for left and right teat numbers, likely due to asymmetry. In our study, excluding individuals without recorded teat numbers, 68 samples exhibited asymmetry between the left and right teat numbers. Furthermore, some samples lacked records for one side of the teats during collection, contributing to discrepancies in heritability estimates.

Conclusion

In summary, we identified 429 CNVRs in the French Yorkshire pig population, covering approximately 2.95% of the total *Sus Scrofa* 11.1 autosomal genome length. These findings complement the CNV map of the Yorkshire pig genome. Our GWAS results revealed 32 variants and three CNVRs significantly associated with teat number. Two critical candidate genes, *TRIM66* and *PRICKLE1*, were related to teat number. The combination of GWAS and multiple genetic mutations presents a valuable approach for enhancing genetics and analyzing the genetic mechanisms underlying pig breeding traits, and contributes to the field of genetic improvement in pig breeding.

Materials and methods

Ethics statement

The animals and experimental procedures used in this study follow the guidelines of the Animal Care and Use Committee of the South China Agricultural University (SCAU) (Guangzhou, China). The ethics committee of SCAU approved all animal experiments (SCAU#2014–0136). The experimental animals were not anesthetized or euthanized in this study. We confirmed that all methods are reported in accordance with ARRIVE guidelines (<https://arriveguidelines.org>) for the reporting of animal experiments.

Animals and phenotype

In this study, a total of 659 French Yorkshire sows were raised in four nucleus pig breeding farms of the Wens Foodstuff Group Co., Ltd. (Guangdong, China) between 2012 and 2016: Yuhe Farm 1 (YH1), Yuhe Farm 2 (YH2), Qingyuan Farm 2 (QY2), and Baizi Farm (BZ). All pigs were subjected to the normal management conditions. Additionally, the left teat number (LTN) and right teat number (RTN) was counted separately after birth, and total teat number (TTN) was the sum of teat numbers on both sides in accordance with our previous study [7].

SNP genotyping and quality control

DNA of each pig was extracted from ear tissue following the standard protocols. The quality of DNA in all samples (659 DNA samples) was measured by a light absorption ratio (A260/280 and A260/230) and gel electrophoresis. All DNA samples were diluted to a concentration of 50 ng/μL. The samples were genotyped with the GeneSeek Porcine 80K SNP chip, which contains 68,528 SNPs uniformly spanning the pig genome. The genotype quality control was conducted by PLINK v1.9 software [52]. SNPs located on the sex chromosomes or without positional information were excluded and a set

of 62,078 SNPs from 659 high-quality genotyping samples (call rate $\geq 90\%$) were retained for CNV detection. Furthermore, to improve the accuracy of the GWAS results, variants with call rates $< 90\%$, minor allele frequency $< 5\%$, and P value $< 10^{-6}$ for the Hardy–Weinberg equilibrium test were also excluded, and individuals with call rates larger than 95% were retained. After quality control, we enhanced the genotype data to the whole-genome sequence level using an imputation strategy. We employed the Swine Imputation (SWIM) Server tool [53] with default parameter settings to perform genotype imputation, bridging the target and reference genotype data. The reference haplotype panels were constructed from whole-genome sequencing data collected from 2259 pigs, representing 44 breeds. The genotype imputation accuracy consistently demonstrated a high average concordance rate exceeding 97%, a non-reference concordance rate of 91%, and an r^2 value of 0.89. This ensured the reliability and robustness of our imputed data. We applied the same quality control criteria to the imputed data as we did to the 80K SNP chip data. After quality control, the final 50294 SNPs and 14656673 variants (including SNPs and INDELs) from 644 French Yorkshire sows were retained for subsequent analysis in 80K chip and SWIM imputed data, respectively.

CNVRs detection and functional enrichment analysis

The PennCNV software v1.0.5 [20] was utilized to identify CNVs by incorporating the SNP signal data of log R ratio (LRR) and B allele frequency (BAF) for each individual. The CNV calling and CNVR determined were carried out following our previous study [35]. In brief, raw CNV dataset, poor quality samples ($n=11$) were filtered out with the following criteria: LRR > 0.3 , BAF drift > 0.01 , GC wave factor of LRR > 0.05 , and then CNVs with consecutive SNPs ≥ 3 and length ≥ 10 kb were filtered to obtain more reliable CNV calls. Afterward, both BEDTools software v2.26.0 [54] and CNVRuler software v1.3.3.2 [55] were used to merge CNVs with at least 1 bp overlap in all samples to determine the CNVR [31]. Finally, 8,746 CNVs and 429 CNVRs were identified. KOBAS v3.0 [56] was used for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of genes which involved in CNVR map. In the enrichment analysis, the statistical method of Fisher's exact test was used to retain the GO entries and pathways with P value < 0.05 . Besides, the CNVRs were mapped to pig QTL from the Animal QTL database (<https://www.animalgenome.org/cgi-bin/QTLdb/index>) [57]. To ensure the accuracy and validity of the GWAS results, we filtered the CNVRs dataset by removing those with frequencies smaller than 1% and 130 CNVRs were retained for GWAS analysis.

GWAS for teat number

To identify candidate variants associated with teat number, GWAS were performed with the CNVR dataset and the SNP dataset, separately.

In this study, GWASs were performed separately using a univariate linear mixed model implemented in the GEMMA software v0.98.1 [58]. Before GWAS, genomic relatedness matrix (GRM) and principal component analysis (PCA) were estimated using the GEMMA and GCTA software v1.92.4beta [59] based on SNP datasets. The statistical model used was as follows:

$$y = W + X + u +$$

where y represents a vector of the phenotypic value for all animals; W is the incidence matrix of covariates, including fixed effects of the top five eigenvectors of PCA; X represents the vector of corresponding coefficients including the intercept; X is the vector of all marker genotypes; u specifies the corresponding effect size of the marker; u is the vector of random effects, with $u \sim MVN_n(0, -^1K)$; σ^2 is the vector of random residuals, with $\sigma^2 \sim MVN_n(0, -^1In)$; σ^2 signifies the ratio between two variance components; σ^2 is the variance of the residual errors; K is GRM; I is an $n \times n$ identity matrix; MVN_n denotes the n -dimensional multivariate normal distribution. In the 80 K chip GWAS, the significance cutoff was defined as the Bonferroni method; a stringent genome-wide threshold (significant) as well as a more lenient chromosome-wide threshold (suggestive) were $P < 0.05/N$ and $P < 1/N$, respectively, where N is the number of variants or CNVRs tested in the analyses. Based on human GWAS results, we set the genome-wide significance threshold and suggestive significance threshold for GWAS based on imputed data at $5.00E-8$ and $1.00E-6$, respectively [60, 61]. Besides, Haploview [62] was used for haplotype block analysis to detect linkage disequilibrium (LD) among SNPs, with settings to "Ignore pairwise comparisons of markers > 500 kb apart" and "Exclude individuals with > 50% missing genotypes".

Candidate genes identification

In this study, the position of SNPs was based on the *Sus Scrofa* 11.1 version of the pig reference genome. We conducted a functional gene annotation to identify candidate genes using Mouse Genome Informatics website (<https://www.informatics.jax.org/>), GeneCards (<http://www.genecards.org/>), GeneCards website (<http://www.genecards.org/>) and Ensembl website (www.ensembl.org/biomart/martview).

Abbreviations

GWAS Genome-wide association analysis
SNP Single nucleotide polymorphism

CNV Copy number variation
CNVR Copy number variation region
TTN Total teat number
LTN Left teat number
RTN Right teat number
SSC *Sus scrofa* Chromosome
QTL Quantitative trait locus

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-024-10611-9>.

Additional file 1: Table S1. Overview of CNVRs for French Yorkshire pigs.

Additional file 2: Table S2. KEGG and GO Enrichment Analysis for all CNVRs gene set by KOBAS.

Additional file 3: Figure S1. Top 10 of enriched KEGG pathways (a) and top 10 of GO terms (b) enrichment, re-spectively.

Additional file 4: Table S3. QTLs included in or partially overlapped with identified CNVRs.

Additional file 5: Table S4. Comparison of CNVRs identified in this study with other studies (based on the *Sscrofa* 11.1 genome assembly).

Additional file 6: Figure S2. PCA plot of population structure based on SNPs. YH1, YH2, QY2, and BZ represent the four pig farms where we collected samples, namely Yuhe Farm 1, Yuhe Farm 2, Qingyuan District 2, and Baizi Farm.

Additional file 7: Figure S3. Q-Q plots showing the observed versus expected log P -values for LTN, RTN and TTN, respectively.

Additional file 8: Figure S4. Haplotype blocks on SSC5. Haplotype blocks are marked with triangles. Values in boxes are the Linkage disequilibrium (r^2) between the SNP pairs. The haplotype blocks are colored in accordance with the mode of standard D'/LOD color scheme.

Additional file 9: Figure S5. Manhattan plots of imputed data GWAS in this population. Manhattan plots consisted of TTN (a), LTN (b) and RTN (c), respectively. The x-axis represents the chromosomes, and the y-axis represents the $-\log_{10}(P\text{-value})$. The solid and dashed lines indicate the genome-wide significance threshold ($P=5.00E-8$) and suggestive significance threshold ($P=1.00E-6$), respectively.

Additional file 10: Table S5. Significant variants associated with teat number in French Yorkshire pigs.

Additional file 11: Figure S6. PCA plot of population structure based on CNVRs. YH1, YH2, QY2, and BZ represent the four pig farms where we collected samples, namely Yuhe Farm 1, Yuhe Farm 2, Qingyuan District 2, and Baizi Farm.

Acknowledgements

The authors would like to thank all the staff of the pig core breeding farms of Wens Foodstuff Group Co., Ltd. (Guangdong, China) for their help of sample collection.

Authors' Contributions

Z.W. and E.Z. conceived and designed the experiment. D.L., Y.Q., F.Z., X.L., S.D., Q.C., and J.Y. collected the samples and recorded the phenotypes. D.L., Y.Q., F.Z., X.L., and S.D. extracted the DNA for genotyping. Y.Q., D.L., and F.Z. analyzed the data. D.L. and Y.Q. visualized the experimental results. D.L. and Y.Q. wrote the manuscript. G.C., J.Y., and E.Z. revised the manuscript for comments. Z.W. contributed to the materials. All authors have read and agreed to the published version of the manuscript.

Funding

This research was supported by the Key Technologies R&D Program of Guangdong Province project (2022B0202090002) and the National Key Research and Development Program of China (2021YFD1301103).

Availability of data and materials

The variation data reported in this paper have been deposited in the Genome Variation Map in National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences and China National Center for Bioinformation, under accession number GVM000796 (<https://bigd.big.ac.cn/gvm/getProjectDetail?Project=GVM000796>).

Declarations

Ethics approval and consent to participate

The animals and experimental methods used in this study follow the guidelines of the Ministry of Agriculture of China and the Use Committee of South China Agricultural University (SCAU). The ethics committee of SCAU (Guangzhou, China) approved all animal experiments. The informed consent was obtained from Wens Foodstuff Group Co., Ltd. (Guangdong, China) for data collection. There was no use of human participants, data, or tissues. We confirmed that all methods are reported in accordance with ARRIVE guidelines (<https://arriveguidelines.org>) for the reporting of animal experiments.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Present Address: College of Animal Science and National Engineering Research Center for Breeding Swine Industry, South China Agricultural University, Guangzhou 510642, China. ²Guangdong Zhongxin Breeding Technology Co., Ltd, Guangzhou 510642, China. ³Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, South China Agricultural University, Guangzhou 510642, China. ⁴Yunfu Subcenter of Guangdong Laboratory for Lingnan Modern Agriculture, Yunfu 527300, China.

Received: 27 December 2023 Accepted: 10 July 2024

Published online: 25 July 2024

References

- Verardo LL, Silva FF, Lopes MS, Madsen O, Bastiaansen JW, Knol EF, Kelly M, Varona L, Lopes PS, Guimarães SE. Revealing new candidate genes for reproductive traits in pigs: combining Bayesian GWAS and functional pathways. *Genet Sel Evol.* 2016;48:9.
- Zhou LS, Zhao WM, Tu F, Wu YH, Ren SW, Fang XM. Physiology and genetics research progress of teat traits in pigs. *Yi Chuan.* 2019;41(5):384–90.
- Arakawa A, Okumura N, Taniguchi M, Hayashi T, Hirose K, Fukawa K, Ito T, Matsumoto T, Uenishi H, Mikawa S. Genome-wide association QTL mapping for teat number in a purebred population of Duroc pigs. *Anim Genet.* 2015;46(5):571–5.
- Rohrer GA, Nonneman DJ. Genetic analysis of teat number in pigs reveals some developmental pathways independent of vertebra number and several loci which only affect a specific side. *Genet Sel Evol.* 2017;49(1):4.
- Li Y, Pu L, Shi L, Gao H, Zhang P, Wang L, Zhao F. Revealing new candidate genes for teat number relevant traits in duroc pigs using genome-wide association studies. *Animals : an open access journal from MDPI.* 2021;11(3):806.
- Yang J, Huang L, Yang M, Fan Y, Li L, Fang S, Deng W, Cui L, Zhang Z, Ai H, et al. Possible introgression of the VRTN mutation increasing vertebral number, carcass length and teat number from Chinese pigs into European pigs. *Sci Rep.* 2016;6: 19240.
- Zhuang Z, Ding R, Peng L, Wu J, Ye Y, Zhou S, Wang X, Quan J, Zheng E, Cai G, et al. Genome-wide association analyses identify known and novel loci for teat number in Duroc pigs using single-locus and multi-locus models. *BMC Genomics.* 2020;21(1):344.
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747–53.
- Du RQ, Jin L, Zhang F. [Copy number variations in the human genome: their mutational mechanisms and roles in diseases]. *Yi Chuan.* 2011;33(8):857–69.
- Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, Horsman DE, MacAulay C, Ng RT, Brown CJ, Eichler EE, et al. A comprehensive analysis of common copy-number variations in the human genome. *Am J Hum Genet.* 2007;80(1):91–104.
- Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shaperro MH, Carson AR, Chen W, et al. Global variation in copy number in the human genome. *Nature.* 2006;444(7118):444–54.
- Zhang F, Gu W, Hurles ME, Lupski JR. Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet.* 2009;10:451–81.
- Iafate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C. Detection of large-scale variation in the human genome. *Nat Genet.* 2004;36(9):949–51.
- Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Månér S, Massa H, Walker M, Chi M, et al. Large-scale copy number polymorphism in the human genome. *Science.* 2004;305(5683):525–8.
- Buckland PR. Polymorphically duplicated genes: their relevance to phenotypic variation in humans. *Ann Med.* 2003;35(5):308–15.
- Yang P, Zhang Z, Xu J, Qu K, Lyv S, Wang X, Cai C, Li Z, Wang E, Xie J, et al. The association of the copy number variation of the MLLT10 gene with growth traits of Chinese cattle. *Animals.* 2020;10(2):250.
- Kang X, Li M, Liu M, Liu S, Pan MG, Wiggins GR, Rosen BD, Liu GE. Copy number variation analysis reveals variants associated with milk production traits in dairy goats. *Genomics.* 2020;112(6):4934–7.
- Chen C, Liu C, Xiong X, Fang S, Yang H, Zhang Z, Ren J, Guo Y, Huang L. Copy number variation in the MSRB3 gene enlarges porcine ear size through a mechanism involving miR-584-5p. *Genet Select Evol.* 2018;50(1):72.
- Wang Y, Zhang T, Wang C. Detection and analysis of genome-wide copy number variation in the pig genome using an 80 K SNP Beadchip. *J Anim Breed Genet.* 2020;137(2):166–76.
- Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, Hakonarson H, Bucan M. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* 2007;17(11):1665–74.
- Chen C, Qiao R, Wei R, Guo Y, Ai H, Ma J, Ren J, Huang L. A comprehensive survey of copy number variation in 18 diverse pig populations and identification of candidate copy number variable genes associated with complex traits. *BMC Genomics.* 2012;13: 733.
- Wang J, Wang H, Jiang J, Kang H, Feng X, Zhang Q, Liu JF. Identification of genome-wide copy number variations among diverse pig breeds using SNP genotyping arrays. *PLoS One.* 2013;8(7): e68683.
- Fernández AI, Barragán C, Fernández A, Rodríguez MC, Villanueva B. Copy number variants in a highly inbred Iberian porcine strain. *Anim Genet.* 2014;45(3):357–66.
- Wang Y, Tang Z, Sun Y, Wang H, Wang C, Yu S, Liu J, Zhang Y, Fan B, Li K, et al. Analysis of genome-wide copy number variations in Chinese indigenous and western pig breeds by 60 K SNP genotyping arrays. *PLoS One.* 2014;9(9): e106780.
- Wiedmann RT, Nonneman DJ, Rohrer GA. Genome-wide copy number variations using SNP genotyping in a mixed breed swine population. *PLoS One.* 2015;10(7): e0133529.
- Xie J, Li R, Li S, Ran X, Wang J, Jiang J, Zhao P. Identification of copy number variations in Xiang and Kele Pigs. *PLoS One.* 2016;11(2): e0148565.
- Wang J, Jiang J, Wang H, Kang H, Zhang Q, Liu JF. Improved detection and characterization of copy number variations among diverse pig breeds by array CGH. *G3 (Bethesda).* 2015;5(6):1253–61.
- Stafuzza NB, Silva RMO, Fragomeni BO, Masuda Y, Huang Y, Gray K, Lourenco DAL. A genome-wide single nucleotide polymorphism and copy number variation analysis for number of piglets born alive. *BMC Genomics.* 2019;20(1):321.
- Wang C, Chen H, Wang X, Wu Z, Liu W, Guo Y, Ren J, Ding N. Identification of copy number variations using high density whole-genome SNP markers in Chinese Dongxiang spotted pigs. *Asian Australas J Anim Sci.* 2019;32(12):1809–15.
- Keel BN, Nonneman DJ, Lindholm-Perry AK, Oliver WT, Rohrer GA. A survey of copy number variation in the porcine genome detected from whole-genome sequence. *Front Genet.* 2019;10: 737.

31. Zheng X, Zhao P, Yang K, Ning C, Wang H, Zhou L, Liu J. CNV analysis of Meishan pig by next-generation sequencing and effects of AHR gene CNV on pig reproductive traits. *Journal of animal science and biotechnology*. 2020;11:42.
32. Bovo S, Ribani A, Muñoz M, Alves E, Araujo JP, Bozzi R, Charneca R, Di Palma F, Etherington G, Fernandez AI, et al. Genome-wide detection of copy number variants in European autochthonous and commercial pig breeds by whole-genome sequencing of DNA pools identified breed-characterising copy number states. *Anim Genet*. 2020;51(4):541–56.
33. Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA*. 2008;299(11):1335–44.
34. Brewer C, Holloway S, Zawalnyski P, Schinzel A, FitzPatrick D. A chromosomal duplication map of malformations: regions of suspected haplo- and triplolethality—and tolerance of segmental aneuploidy—in humans. *Am J Hum Genet*. 1999;64(6):1702–8.
35. Qiu Y, Ding R, Zhuang Z, Wu J, Yang M, Zhou S, Ye Y, Geng Q, Xu Z, Huang S, et al. Genome-wide detection of CNV regions and their potential association with growth and fatness traits in Duroc pigs. *BMC Genomics*. 2021;22(1):332.
36. Lu W, Zhang Y, Liu D, Songyang Z, Wan M. Telomeres-structure, function, and regulation. *Exp Cell Res*. 2013;319(2):133–41.
37. Jiang L, Jiang J, Yang J, Liu X, Wang J, Wang H, Ding X, Liu J, Zhang Q. Genome-wide detection of copy number variations using high-density SNP genotyping platforms in Holsteins. *BMC Genomics*. 2013;14: 131.
38. Zhang D, Qian Y, Akula N, Alliey-Rodriguez N, Tang J, Gershon ES, Liu C. Accuracy of CNV detection from GWAS data. *PLoS One*. 2011;6(1): e14511.
39. Jiang J, Wang J, Wang H, Zhang Y, Kang H, Feng X, Wang J, Yin Z, Bao W, Zhang Q, et al. Global copy number analyses by next generation sequencing provide insight into pig genome variation. *BMC Genomics*. 2014;15(1): 593.
40. Liu GE, Hou Y, Zhu B, Cardone MF, Jiang L, Cellamare A, Mitra A, Alexander LJ, Coutinho LL, Dell'Aquila ME, et al. Analysis of copy number variations among diverse cattle breeds. *Genome Res*. 2010;20(5):693–703.
41. Steemers FJ, Gunderson KL. Whole genome genotyping technologies on the BeadArray platform. *Biotechnol J*. 2007;2(1):41–9.
42. Ning L, Huo Q, Xie N. Comprehensive analysis of the expression and prognosis for tripartite motif-containing genes in breast cancer. *Front Genet*. 2022;13: 876325.
43. Zhang H, Zheng Y, Zhang Y. Knockdown of TRIM66 in MDA-MB-468 triple negative breast cancer cell line suppresses proliferation and promotes apoptosis through EGFR signaling. *Pol J Pathol*. 2021;72(2):160–6.
44. Liu Z, Li H, Zhong Z, Jiang S. A whole genome sequencing-based genome-wide association study reveals the potential associations of teat number in Qingping Pigs. *Animals*. 2022;12(9):1057.
45. Liu C, Lin C, Gao C, May-Simera H, Swaroop A, Li T. Null and hypomorph Prickle1 alleles in mice phenocopy human Robinow syndrome and disrupt signaling downstream of Wnt5a. *Biol Open*. 2014;3(9):861–70.
46. Roarty K, Serra R. Wnt5a is required for proper mammary gland development and TGF-beta-mediated inhibition of ductal growth. *Development (Cambridge, England)*. 2007;134(21):3929–39.
47. Peng WF, Xu SS, Ren X, Lv FH, Xie XL, Zhao YX, Zhang M, Shen ZQ, Ren YL, Gao L, et al. A genome-wide association study reveals candidate genes for the supernumerary nipple phenotype in sheep (*Ovis aries*). *Anim Genet*. 2017;48(5):570–9.
48. Tang J, Zhang Z, Yang B, Guo Y, Ai H, Long Y, Su Y, Cui L, Zhou L, Wang X, et al. Identification of loci affecting teat number by genome-wide association studies on three pig populations. *Asian Australas J Anim Sci*. 2017;30(1):1–7.
49. Lopes MS, Bastiaansen JW, Harlizius B, Knol EF, Bovenhuis H. A genome-wide association study reveals dominance effects on number of teats in pigs. *PLoS ONE*. 2014;9(8): e105867.
50. Wang L, Zhang Y, Zhang T, Zhang L, Yan H, Liu X, Wang L. Genotyping by sequencing reveals a new locus for pig teat number. *Anim Genet*. 2017;48(4):470–2.
51. Wei C, Cai X, Diao S, Teng J, Xu Z, Zhang W, Zeng H, Zhong Z, Wu X, Gao Y, et al. Integrating genome-wide association study with multi-tissue transcriptome analysis provides insights into the genetic architecture of teat traits in pigs. *J Genet Genomics*. 2023;50(10):795–8.
52. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*. 2015;4:7.
53. Ding R, Savegnago R, Liu J, Long N, Tan C, Cai G, Zhuang Z, Wu J, Yang M, Qiu Y, et al. The SWine IMputation (SWIM) haplotype reference panel enables nucleotide resolution genetic mapping in pigs. *Commun Biol*. 2023;6(1):577.
54. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841–2.
55. Kim JH, Hu HJ, Yim SH, Bae JS, Kim SY, Chung YJ. CNVRuler: a copy number variation-based case-control association analysis tool. *Bioinformatics*. 2012;28(13):1790–2.
56. Bu D, Luo H, Huo P, Wang Z, Zhang S, He Z, Wu Y, Zhao L, Liu J, Guo J, et al. KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res*. 2021;49(W1):W317–25.
57. Hu ZL, Park CA, Reecy JM. Bringing the animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. *Nucleic Acids Res*. 2022;50(D1):D956–d961.
58. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet*. 2012;44(7):821–4.
59. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88(1):76–82.
60. Johnson RC, Nelson GW, Troyer JL, Lautenberger JA, Kessing BD, Winkler CA, O'Brien SJ. Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics*. 2010;11: 724.
61. Huang Y, Cai L, Duan Y, Zeng Q, He M, Wu Z, Zou X, Zhou M, Zhang Z, Xiao S, et al. Whole-genome sequence-based association analyses on an eight-breed crossed heterogeneous stock of pigs reveal the genetic basis of skeletal muscle fiber characteristics. *Meat Sci*. 2022;194: 108974.
62. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21(2):263–5.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Article

Increased Accuracy of Genomic Prediction Using Preselected SNPs from GWAS with Imputed Whole-Genome Sequence Data in Pigs

Yiyi Liu ^{1,2,†}, Yuling Zhang ^{1,2,†}, Fuchen Zhou ^{1,2}, Zekai Yao ^{1,2}, Yuexin Zhan ^{1,2}, Zhenfei Fan ^{1,2}, Xianglun Meng ^{1,2}, Zebin Zhang ^{1,2}, Langqing Liu ^{1,2}, Jie Yang ^{1,2}, Zhenfang Wu ^{1,2,3}, Gengyuan Cai ^{1,2,3,*} and Enqin Zheng ^{1,2,*}

¹ National Engineering Research Center for Breeding Swine Industry, College of Animal Science, South China Agricultural University, Guangzhou 510642, China; yiyiliu0921@outlook.com (Y.L.); yulingzhang@stu.scau.edu.cn (Y.Z.); zfc17854225519@163.com (F.Z.); 14770011284@163.com (Z.Y.); zhanyx0923@outlook.com (Y.Z.); fanzhenfei3@gmail.com (Z.F.); a102818222105979@163.com (X.M.); zbzhang@scau.edu.cn (Z.Z.); langqing.liu@scau.edu.cn (L.L.); jieyang2012@hotmail.com (J.Y.); wzfemail@163.com (Z.W.)

² Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, South China Agricultural University, Guangzhou 510642, China

³ Guangdong Zhongxin Breeding Technology Co., Ltd., Guangzhou 510642, China

* Correspondence: cgy0415@163.com (G.C.); eqzheng@scau.edu.cn (E.Z.)

[†] These authors contributed equally to this work.

Simple Summary: By integrating prior biological information into genomic selection methods using appropriate models, it is possible to improve prediction accuracy for complex traits. In this context, we conducted a comparative assessment of two genomic prediction models, namely, genomic best linear unbiased prediction and genomic feature best linear unbiased prediction. The accuracy of these models in predicting the growth traits of backfat thickness and loin muscle area was evaluated. Our results revealed that the genomic feature best linear unbiased prediction model can effectively integrate prior information into the model, which is superior to the genomic best linear unbiased prediction model in some cases. These findings provide valuable ideas for enhancing the genomic prediction accuracy of growth traits in pigs.

Abstract: Enhancing the accuracy of genomic prediction is a key goal in genomic selection (GS) research. Integrating prior biological information into GS methods using appropriate models can improve prediction accuracy for complex traits. Genome-wide association study (GWAS) is widely utilized to identify potential candidate loci associated with complex traits in livestock and poultry, offering essential genomic insights. In this study, a GWAS was conducted on 685 Duroc × Landrace × Yorkshire (DLY) pigs to extract significant single-nucleotide polymorphisms (SNPs) as genomic features. We compared two GS models, genomic best linear unbiased prediction (GBLUP) and genomic feature BLUP (GFBLUP), by using imputed whole-genome sequencing (WGS) data on 651 Yorkshire pigs. The results revealed that the GBLUP model achieved prediction accuracies of 0.499 for backfat thickness (BFT) and 0.423 for loin muscle area (LMA). By applying the GFBLUP model with GWAS-based SNP preselection, the average prediction accuracies for BFT and LMA traits reached 0.491 and 0.440, respectively. Specifically, the GFBLUP model displayed a 4.8% enhancement in predicting LMA compared to the GBLUP model. These findings suggest that, in certain scenarios, the GFBLUP model may offer superior genomic prediction accuracy when compared to the GBLUP model, underscoring the potential value of incorporating genomic features to refine GS models.

Keywords: pigs; GS; accuracy; imputed WGS data; genome-wide association study; SNP preselection

Citation: Liu, Y.; Zhang, Y.; Zhou, F.; Yao, Z.; Zhan, Y.; Fan, Z.; Meng, X.; Zhang, Z.; Liu, L.; Yang, J.; et al.

Increased Accuracy of Genomic Prediction Using Preselected SNPs from GWAS with Imputed Whole-Genome Sequence Data in Pigs. *Animals* **2023**, *13*, 3871.

<https://doi.org/10.3390/ani13243871>

Academic Editor: Martino Cassandro

Received: 2 November 2023

Revised: 13 December 2023

Accepted: 14 December 2023

Published: 15 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Pork accounts for a large share of total global meat production, addressing the increasing demand for high-quality protein. Genetic factors play a predominant role among the various elements influencing the efficiency of the swine industry. It is essential to comprehend and optimize the genetic potential of pigs through genomic selection methods to enhance the efficiency and sustainability of pork production [1]. Genomic selection (GS), first proposed in 2001, is a statistical method for calculating the genomic estimated breeding value (GEBV) using high-density single-nucleotide polymorphisms (SNPs) across the whole genome [2]. This method hypothesizes that quantitative trait loci (QTL) for all target traits are in linkage disequilibrium (L) with at least one marker in the genome-wide high-density single-nucleotide polymorphism (SNP), so that the effect of each QTL can be reflected by the SNP [3]. The GEBV, calculated from individual genome data, outperforms the traditional estimated breeding value (EBV) based on pedigree records. This enhanced accuracy stems from GS, which leverages comprehensive genome-wide genetic marker information to more accurately depict the genetic relatedness among individuals [4]. Since GEBV can be independent of pedigree and phenotypic records, this facilitates selection early, which can significantly shorten the generation interval and increase the accuracy of predicted breeding values [5]. In pig breeding systems, the generation interval of pigs has been controlled for a short time and it is difficult to scale it down [6]. The GS of pigs is primarily based on improving the accuracy of GEBV to obtain additional genetic progression, and its accuracy depends largely on GS approaches.

The two major challenges of GS are the accuracy of GEBV and the cost of genotypes. SNP chip data have largely been the foundation for the adoption of GS throughout the past ten years. With the improvement of the reference genome sequence of livestock and the reduction in the cost of sequencing, whole-genome sequencing (WGS) data have become possible for GS. WGS data would include causal mutations that can find many QTL closely linked to targeting traits, which can greatly improve the accuracy of genomic prediction [7,8]. A study used simulation data to compare the accuracy of GEBV with a GBLUP model under low-density chip data (7.5K SNPs), high-density chip data (17K), and WGS data (335K), and found that the accuracy of GEBV based on WGS data was 4% higher than that based on chip data [9]. While the cost of WGS is falling rapidly, sequencing a large number of animals remains expensive. For most individuals, SNP high-density chips were used for genotyping, and the genetic variation in the whole genome obtained through genotype imputation would be cost-effective [10]. Currently, there are many livestock and poultry that perform GS based on imputed high-density genetic data or imputed sequence data, such as chickens [11,12], pigs [13], and cattle [14,15].

However, it was previously found that, compared to the GS based on chip data, using imputed WGS data for the accuracy of genomic prediction produced no advantage [9,15,16]. One of the important reasons for this is that prior information lacks consideration. At present, there are several ways to use prior biological information for GS. A common one is the use of QTL information, such as GS using QTL information in Holstein dairy cows, which improves accuracy by 4% [17]. Genome-wide association study (GWAS) is based on LD, identifying marker loci closely correlated with phenotypic variation by comparing the relationship between phenotypic differences in target traits across different individuals and the polymorphisms at genetic loci. Given the significance of GWAS in identifying candidate loci for complex traits in livestock and poultry, the selection of SNPs with significant effects on target traits based on GWAS results has become one of the most widely applied genomic-level priors [18,19]. Gebreyesus improved the accuracy of prediction through integrating prior information from GWAS results by 13–38% compared with the genomic best linear unbiased prediction (GBLUP) model in the study of milk fatty acid composition [20]. Subsequent studies also showed that significant SNPs found by GWAS using imputed WGS data can increase the accuracy of genomic prediction [8,10,21]. The genomic feature best linear unbiased prediction (GFBLUP) model, proposed by Edwards [22], can integrate biological prior information and treat

markers that significantly impact traits as independent random effects within the model. This approach has been shown to enhance the accuracy of genomic prediction through the integration of prior information.

Duroc × Landrace × Yorkshire (DLY) hybrid pigs are the most widely bred pigs in the swine industry, as they have the advantages of fast growth and high feed utilization, providing consumers with the largest source of pork [23]. In this work, we adopt a research strategy of integrating prior information from GWAS results to evaluate the genomic prediction accuracy of the GBLUP and GFBUP models using imputed WGS data. We collected important growth traits of pigs, backfat thickness (BFT) and loin muscle area (LMA). The objectives of this study were (i) to perform a GWAS on a total of 685 Duroc × Landrace × Yorkshire (DLY) pigs to extract significant SNPs as genomic prior information; (ii) to improve the accuracy of genomic prediction by integrating preselected SNPs from the GWAS into the prediction models using a population of 651 Yorkshire pigs.

2. Materials and Methods

2.1. Ethics Statement

All animals used in this study met the guidelines for the care and use of experimental animals established by the Ministry of Agriculture of China. The whole study was approved by the ethics committee of South China Agriculture University (SCAU, Guangzhou, China). The experimental animals were not anesthetized or euthanized in order to conduct this study.

2.2. Pig Population

The experimental animals in this study were two populations, the prior discovery population and the reference/validation population. A total of 685 DLY pigs (338 males and 347 females) were provided by Wens Foodstuffs Group Co., Ltd. (Yunfu, Guangdong, China) and served as the prior discovery population, born from May to August 2019. DLY pigs were bred from Duroc boars, including American Duroc (S21) pigs, Canadian Duroc (S22) pigs, and Taiwanese Duroc (S23) pigs, crossed with Landrace × Yorkshire sows. In total, 651 Yorkshire pigs provided by Guangdong Guangken Animal Husbandry Group Co., Ltd., born between July 2019 and October 2020, were used as the reference and validation populations. All pigs were raised with the same customized diet in human-controlled farm conditions and similar management conditions. The customized corn–soybean feed (free of probiotics and antibiotics) contained 16% crude protein, 3100 kJ of digestible energy, and 0.78% lysine. Water was available *ad libitum*. Feeding was completed when the body weight reached 100 ± 5 kg.

2.3. Phenotypic Data

According to the current effective standard, large-scale pig slaughtering experiments were carried out. Ear tissue was collected as follows: the pig's ear was first cleaned with 75% alcohol. Then, a clear forfex was used to cut out a small fraction of ear tissue. The wound was then treated with tincture of iodine. The protocol for ear tissue collection was approved by the ethics committee of SCAU. The main growth traits of pigs were collected, that is, back fat thickness (BFT) and loin muscle area (LMA). BFT was collected with vernier caliper at rib 6–7 thoracic vertebra when pigs weighed 100 ± 5 kg, and the maximum height and width of the cross-sectional area of the longissimus dorsi muscle at the last rib were measured. LMA is calculated as follows:

$$LMA(cm^2) = height \times width \times 0.7$$

After collecting phenotypic data, the original data underwent a correction process using the single-trait animal model in the BLUPF90 programs [24]. The model used for correction included fixed effects such as sex, age, and the first three principal components. The MVP package in R language was used to analyze the phenotypic distribution of the

reference/validation population. The resulting corrected phenotypic values were used for all subsequent analyses.

2.4. Genotype Data

The genomic DNA was isolated and extracted from approximately 15–20 mg of ear tissue following the traditional phenol/chloroform method. The quality and quantity of the DNA samples were measured with a NanoDrop™ 2000 (Thermo Fisher Scientific, Waltham, MA, USA) as previously described [25]. In total, 685 DLY pigs were genotyped using the GeneSeek Porcine 50K SNP Chip (Neogen, Lincoln, NE, USA). The original genotype data can be read using the GenomeStudio 2.0.5 software. The genotype quality control procedures were performed using PLINK v1.9 software [26] with the following criteria: (1) individual call rate > 90%; (2) SNP call rate > 90%; (3) minor allele frequency > 1%; (4) Hardy–Weinberg test p -value > 10^{-6} ; and (5) SNPs lacking informative data and those located on the sex chromosomes were removed. After quality control, a final dataset of 33,197 SNPs remained for DLY pigs. Similarly, 651 Yorkshire pigs were genotyped using the GeneSeek Porcine 80K SNP Chip, and for quality control. Due to the different chip types and densities used in these two populations, and in order to increase the density of genetic markers, we performed genotype imputation.

The imputation process of 50K SNP chip data and 80K SNP chip data to WGS genotypes was performed with the SWIM database. The SWIM pig haplotype reference panel, based on 2259 animals across 44 breeds, demonstrated robust performance in genotype imputation, achieving a concordance rate above 96% and an r^2 of 0.85 [27]. Subsequently, the imputed WGS data were quality-controlled using the same standard as above. In the prior discovery population, 15,743,104 SNPs for DLY pigs remained after quality control. For the 651 Yorkshire pigs, 14,166,374 SNPs remained after quality control. The imputed WGS data were used for subsequent analysis.

2.5. Genetic Parameter Estimation

Heritability (h^2) was defined as the ratio of the additive genetic variance to phenotypic variance. Firstly, the genomic relationship matrix (GRM) was constructed to assess kinship among individuals. Subsequently, variance components were estimated using the restricted maximum likelihood (REML) algorithm via the GCTA 1.93.2 software [28].

2.6. Genome-Wide Association Study

For the two traits analyzed, the 685 DLY pigs were used to perform a GWAS and a genomic relationship matrix (GRM) based on all SNPs, which was constructed to account for population structure. The mixed linear model-based association analysis (MLMA) in the package GCTA [29] was used. The GWAS model was as follows:

$$y = 1\mu + Zg + bx + e$$

where y is the vector of corrected phenotypic value; μ is the overall mean, 1 is a vector of ones; g is the random effect, i.e., the accumulated effect of all SNPs, following a normal distribution $g \sim N(0, G\sigma_g^2)$, where G is the GRM which is built using imputed WGS data, captured genetic relatedness, σ_g^2 is the additive genetic variance, Z is the incidence matrix for g ; b is the fixed effect of the candidate SNP to be tested for association, x is a vector of the SNPs' genotype indicator variable coded as 0, 1, or 2; e is a vector of random residuals with $e \sim N(0, I\sigma_e^2)$, where σ_e^2 is the residual variance and I is an identity matrix. For ease of computation, σ_g^2 is estimated based on the null model and then fixed while testing for the association between each SNP and the trait.

2.7. SNP Preselection Based on the GWAS Results

After GWAS analysis, SNPs associated with BFT and LMA traits were categorized into different classes based on their p -values, specifically, those less than 0.05, 0.005, 0.0005, and 0.00005. Each trait resulted in four distinct sets of prior SNP information. Subsequently, these four sets of prior SNP information were individually intersected with the imputed WGS data of the 651 Yorkshire pigs. The common SNPs that are present in both the prior SNP information and the imputed WGS data of 651 Yorkshire pigs served as genomic features.

2.8. Genomic Prediction Models

The statistical methods used for predicting breeding value in this study included genomic best linear unbiased prediction (GBLUP) and genomic feature best linear unbiased prediction (GFBLUP) models. The genotype datasets for the GBLUP and GFBLUP models are imputed WGS data and preselected imputed WGS SNP data based on GWAS results. The GBLUP model is used to calculate GEBV as follows:

$$y = 1\mu + Zu + e$$

where y is the vector of the corrected phenotypic value; 1 is a vector of ones, μ is the overall mean; u is the vector of additive genetic values, and it is assumed that $u \sim N(0, G\sigma_u^2)$, where G is a relationship matrix built with the HIBLUP 1.1.0 software [30]; σ_u^2 represents corresponding additive genetic variance; Z is incidence matrices relating the additive genetic values to the phenotype value; e is the vector of random residual effect, and it is assumed that $e \sim N(0, I\sigma_e^2)$, where σ_e^2 is the residual variance and I is an identity matrix.

The GFBLUP model was an extended BLUP including two random genetic effects:

$$y = 1\mu + Zf + Zr + e$$

where y , 1 , μ , and e are the same as in GBLUP, f is the vector of genetic effects captured by genomic features of target traits, following a normal distribution $f \sim N(0, G_f\sigma_f^2)$, r is the vector of genomic effects captured by the remaining genetic markers in the imputed WGS data that remove genomic features, following a normal distribution $r \sim N(0, G_r\sigma_r^2)$, and Z is an incidence matrix linking (f and r) to the phenotypic values. The G_f and G_r were constructed using only the genetic marker set defined by the genomic feature and the remaining set of markers; all of this was accomplished using the GCTA software. Since the computational resources of using two G matrices are too high, the two G matrices are combined into one G matrix to predict the GEBV:

$$G_{new} = \lambda G_f + (1 - \lambda)G_r$$

where $\lambda = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_r^2}$, σ_f^2 , and σ_r^2 are the genetic variances explained by G_f and G_r for BFT and LMA traits.

2.9. Prediction Accuracy

Cross-validation is a common method used to evaluate the performance of a model. In this study, five-fold cross-validation was performed to assess the prediction accuracy of the model, which was measured by the correlation coefficient between phenotype values and GEBV. Specifically, 651 Yorkshire individuals were randomly divided into 5

groups, with 4 groups serving as reference groups and the remaining group serving as the validation group. The genotypes and phenotype values of the reference groups were known and used to predict the model, while the genotypes of the validation group were known and the phenotype values were treated as missing values. This process was repeated for each group tested. The above steps were repeated 5 times, the prediction accuracy was reported as the average of these 25 results.

3. Results

3.1. Descriptive Statistics of Phenotypes and Heritability

Phenotypic statistics of BFT and LMA were performed in two populations. As Table 1 shows, the mean and standard deviation of BFT and LMA traits in DLY pigs were 11.49 ± 3.27 and 40.44 ± 7.39 , and the coefficient of variation was 28.45% and 18.28%. It is implied that these two phenotypic variations are relatively stable in this population. Furthermore, the values (standard error) of the heritability estimates were 0.35 ± 0.08 and 0.34 ± 0.08 for BFT and LMA traits, both of which were medium heritability traits. Similarly, as shown in Table 2, the mean and standard deviation of BFT and LMA traits in Yorkshire pigs were 11.85 ± 2.43 and 40.34 ± 4.93 , and the coefficient of variation was 20.54% and 12.21%. Data from Table 2 can be compared with the data in Table 1 which show that Yorkshire pigs' heritability of BFT and LMA traits is 0.45, which is slightly higher than the heritability of BFT (0.35) and LMA (0.34) traits of DLY pigs. After phenotypic correction, BFT and LMA were roughly in line with normal distribution in the Yorkshire population (Figure S1), which could be used for GS.

Table 1. Descriptive statistics and heritability of BFT and LMA traits in DLY pigs.

Trait	Unit	Mean (\pm SD) ³	Min ⁴	Max ⁵	C.V./% ⁶	h ² (\pm SE) ⁷
BFT ¹	mm	11.49 ± 3.27	5.50	25.58	25.48	0.35 ± 0.08
LMA ²	cm ²	40.44 ± 7.39	20.25	64.63	18.28	0.34 ± 0.08

¹ Back fat thickness (BFT). ² Loin muscle area (LMA). ³ Standard deviations (SD). ⁴ Minimum (Min).

⁵ Maximum (Max). ⁶ Coefficient of variation (C.V.). ⁷ Heritability (standard error) value (h² (\pm SE)).

Table 2. Descriptive statistics and heritability of BFT and LMA traits in Yorkshire pigs.

Trait	Unit	Mean (\pm SD) ³	Min ⁴	Max ⁵	C.V./% ⁶	h ² (\pm SE) ⁷
BFT ¹	mm	11.85 ± 2.43	5.93	23.23	20.54	0.45 ± 0.06
LMA ²	cm ²	40.34 ± 4.93	24.25	54.67	12.21	0.45 ± 0.07

¹ Back fat thickness (BFT). ² Loin muscle area (LMA). ³ Standard deviations (SD). ⁴ Minimum (Min).

⁵ Maximum (Max). ⁶ Coefficient of variation (C.V.). ⁷ Heritability (standard error) value (h² (\pm SE)).

3.2. SNP Preselection Based on the GWAS Results

The number of SNPs varies in GWAS results using imputed WGS data based on different levels of *p*-value (0.05, 0.005, 0.0005, 0.00005); see Table 3 for details. Based on the GWAS results, about 517.1K, 47.9K, 2.8K, and 171 significant SNPs were preselected according to different *p*-values for BFT traits in 651 Yorkshire pigs. In LMA traits, about 515.5K, 57.7K, 6.3K, and 341 significant SNPs were preselected according to different *p*-values as genomic features.

Table 3. Number of SNPs preselected based on the GWAS results with different *p*-value cutoffs.

<i>p</i> -Value	BFT		LMA	
	Number of SNPs in GWAS ¹	Number of SNPs in Gf ²	Number of SNPs in GWAS ¹	Number of SNPs in Gf ²
<0.05	731,061	517,173	720,230	515,502
<0.005	66,867	47,987	81,028	57,767
<0.0005	4399	2819	8020	6357
<0.00005	262	171	478	341

¹ The number of SNPs divided according to *p*-value based on the GWAS results of DLY pigs (number of SNPs in GWAS). ² The number of SNPs extracted as genomic features in Yorkshire pigs (number of SNPs in Gf).

3.3. Genomic Prediction

The prediction accuracy values of both GBLUP and GFBLUP models are shown in Table 4. The GBLUP model was run using the complete imputed WGS data, and the prediction accuracy values were 0.499 for the BFT trait and 0.423 for the LMA trait. Using SNP preselection based on the GWAS results with different *p*-value cutoffs, the prediction accuracy values of GFBLUP were 0.488, 0.487, 0.487, and 0.491 for the BFT trait and 0.440, 0.420, 0.417, and 0.423 for the LMA trait. The results showed that, for preselected SNPs from the GWAS results with the optimal *p*-value cutoffs ($p < 0.05$), the highest accuracy of the GFBLUP model was 0.440 for the LMA trait. This prediction accuracy value was still higher than that of the GBLUP model with the imputed WGS data. Using preselected SNPs from the GWAS results with the optimal *p*-value cutoffs, the accuracy of the GFBLUP model was lower than that of the GBLUP model for the BFT trait. Overall, the average prediction accuracy of BFT and LMA traits of the GFBLUP model reached the highest values at 0.491 and 0.440, respectively. The prediction accuracy of the GBLUP and GFBLUP models for BFT and LMA traits was visually compared, as depicted in Figure 1. In terms of BFT traits, there was no obvious trend for the accuracy of the GFBLUP model using different *p*-value cutoffs to preselect SNPs. However, for the LMA trait, the GFBLUP model improved prediction accuracy by 4.8% compared to the GBLUP model. In general, using the SNPs preselected from imputed WGS data based on GWAS results led to greater accuracy of genomic prediction for the LMA trait.

Table 4. Prediction accuracies of the GBLUP model and the GFBLUP model with SNP preselection based on GWAS results.

Model	<i>p</i> -Value	Accuracy (Mean ± SE ⁴)	
		BFT	LMA
GBLUP ¹	All ³	0.499 ± 0.016	0.423 ± 0.010
GFBLUP ²	<0.05	0.488 ± 0.017	0.440 ± 0.011
	<0.005	0.487 ± 0.017	0.420 ± 0.011
	<0.0005	0.487 ± 0.016	0.417 ± 0.010
	<0.00005	0.491 ± 0.016	0.423 ± 0.010

¹ Genomic best linear unbiased prediction (GBLUP). ² Genomic feature best linear unbiased prediction (GFBLUP). ³ All SNPs of imputed WGS data (All). ⁴ Standard error (SE).

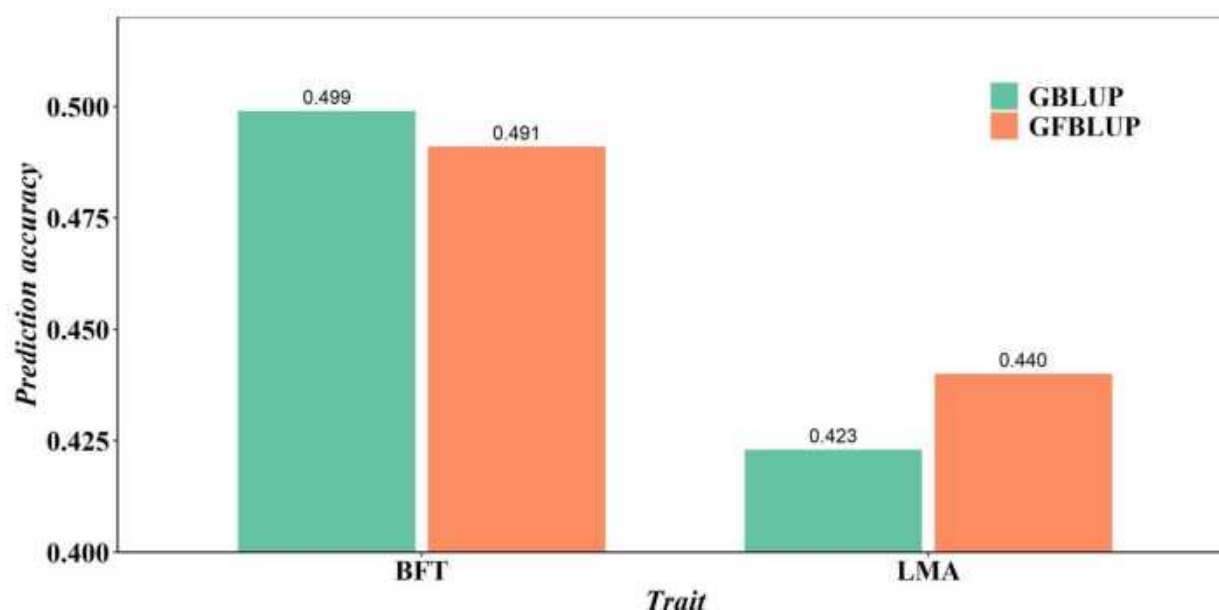


Figure 1. Accuracy of genomic predictions using GBLUP and GFBLUP models for BFT and LMA traits.

4. Discussion

In this study, an independent prior discovery population was specifically designed for GWAS analysis to select prior information. During the initial phases of GS research that integrated prior biological information, the availability of such prior information was relatively limited, often leading to a reliance on reference populations. Nevertheless, bias can be amplified when the prior discovery population and the reference population are the same or when GS relies solely on prior information [15,31,32]. For instance, in cattle research, utilizing the same prior discovery population for a GWAS as the reference population has been demonstrated to result in significant biases in GS [18]. In our research, in total, 685 DLY pigs were used as the prior discovery population for GWAS analysis, integrating completely independent GWAS results, and preselected significant SNPs were based on an independent population. In addition, a separate group of 651 Yorkshire pigs was used as the reference and validation populations to evaluate the prediction accuracy of the GS model. These two populations were relatively independent of each other (Figure S2).

Based on the GWAS results for BFT and LMA traits, significant SNPs were preselected for each trait and incorporated into the imputed WGS data for GS of the respective trait. Consequently, the number and specific sites of significant SNPs differed for each trait. The notable advantage of this study lies in the uniqueness of significant SNPs for each trait, potentially resulting in heightened prediction accuracy. However, it is worth noting that this study relies on prior GWAS information specific to individual traits. Therefore, in situations where a particular trait is not measured within the population or is challenging to obtain, the approach may not be universally applicable. Nonetheless, it is essential to acknowledge that this study relies on prior GWAS information tailored to individual traits. Thus, in cases where a specific trait is unmeasured within the prior discovery population or is challenging to acquire, the approach may not be universally applicable.

We explored the influence of different p -value cutoffs on preselected SNPs in genomic prediction, finding that the optimal p -value threshold significantly impacted prediction accuracy, though no definitive trend emerged across all cases. The preselected SNPs from the GWAS results were categorized based on different levels of p -values (0.05, 0.005, 0.0005, 0.00005). Similar categorization of preselected SNPs has also been utilized in

previous studies [33]. While the categorization method may involve subjectivity, it is important to note that there is currently no strict standard in place. When the p -value is less than 0.05, the majority of potential genetic variants were selected. As the selection criteria become increasingly stringent, such as $p < 0.00005$, only a few hundred genetic variants remain as prior information. The outcomes of this study demonstrated that when using preselected SNPs from the GWAS results with the optimal p -value cutoffs ($p < 0.05$), the highest accuracy of GFBLUP was 0.440 for the LMA trait (Table 4). On the other hand, using preselected SNPs from the GWAS results with a p -value less than 0.00005 resulted in an accuracy of 0.491 for the BFT trait. It is worth noting that a study has speculated that as the number of preselected SNPs increases, the prediction accuracy of GS should first rise and then decline [15]. However, there was no distinct trend in the accuracy of GFBLUP using various p -value cutoffs for preselecting SNPs.

Incorporating prior biological information into GS has been shown to enhance the prediction accuracy for complex traits [34]. Presently, GS integrating prior biological information is widespread in the research focusing on important economic traits in cattle. In milk fatty acid-related traits, the prediction accuracy of GFBLUP integrated with GWAS results increased by an average of 23% in the Danish dairy cattle population and 13% in the Chinese population [20]. Regarding carcass traits in Hanwoo beef cattle, it was found that compared with the prediction accuracy using a 50K benchmark chip, using preselected SNPs from GWAS improved accuracy of prediction by 2.0% to 5.0% [32]. In a sheep study, the accuracy of GS for six meat traits and two wool traits was improved by integrating prior information based on GWAS with the foundation of the 50K chip [35]. A fundamental question revolves around the utilization of this prior information, especially in relation to selecting an appropriate GS model. In this study, we employed the GFBLUP model to integrate preselected significant SNPs. The widely used GBLUP model assumes that the influence of each genomic locus on the trait is uniform, thereby limiting its ability to integrate biological prior information. In contrast, the GFBLUP model overcomes this limitation by fitting two G matrices to assign different genetic weights to different classes of SNPs. Furthermore, the GFBLUP model allows for the fitting of more G matrices and further weighting of G matrices, by weighting multiple G matrices into a single G matrix to reduce computational resources, which can reduce computation time and memory. Consequently, this approach not only eases the memory demands for datasets featuring large sample sizes in the study but also enables the seamless integration of biological prior information into the modeling process. Overall, in terms of application scenarios, the GFBLUP model is more suitable for integrating prior information or other multi-omics data for prediction, as it can incorporate two or more random effects.

A comparison of the prediction accuracy of BFT and LMA traits by the GBLUP and GFBLUP models is depicted in Figure 1. For the GFBLUP model using preselected SNPs from the GWAS results with the different p -value cutoffs, the highest accuracy was 0.440 for the LMA trait and 0.491 for the BFT trait. The GBLUP model with complete imputed WGS data yielded accuracies of 0.423 for the LMA trait and 0.499 for the BFT trait. In the context of predicting the LMA trait in pigs, the GFBLUP model resulted in a notable 4.8% enhancement in genomic prediction accuracy compared to the GBLUP model. Previous research has also indicated that the GFBLUP model yielded a 38% enhancement in genomic prediction accuracy for bovine fatty acid traits [20]. Unfortunately, the prediction accuracy of the GFBLUP model did not meet expectations in predicting the BFT trait. This could be because the proportion of QTL in preselected genomic feature markers was very low, and using preselected SNPs would not be advantageous [36]. Similar results were also found in the *Drosophila* research [33] in terms of the starvation resistance trait. In a study of aquaculture species, it was observed that, compared with the prediction accuracy of the GBLUP model, the accuracy of the GFBLUP model integrating the GWAS results was reduced by an average of 6.2% in the prediction of disease resistance and growth traits [37]. If the proportion of QTL in preselected genomic feature markers was large, the GFBLUP model further increases its prediction accuracy compared to GBLUP with the

complete WGS data [22]. With the same method, the differences in accuracy of genomic prediction between different traits may be due to the underlying genetic architecture [38]. Our future efforts will focus on using multi-omics data to select feature markers, ultimately improving genomic prediction accuracy.

5. Conclusions

Based on the GWAS results of 685 DLY pigs as prior information, we compared the GBLUP and GFBLUP models for accuracy of genomic prediction of two traits (BFT and LMA) by using the imputed WGS data of 651 Yorkshire pigs. Our results revealed that the average genomic prediction accuracy of the GFBLUP model for LMA trait was 4.8% higher than that of the GBLUP model. However, when predicting the BFT trait, both models exhibited comparable levels of prediction accuracy. This suggests that the GFBLUP model can effectively integrate prior information into the model, which is superior to the GBLUP model in some cases. Our findings provide valuable ideas for improving the genomic prediction accuracy of growth traits in pigs.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/ani13243871/s1>. Figure S1: Corrected phenotypic distribution of 651 Yorkshire pigs. (a) Phenotypic distribution of backfat thickness (BFT) trait and (b) Phenotypic distribution of loin muscle area (LMA) trait; Figure S2: Genetic structure of two populations (DLY and Yorkshire pigs), scatter plots of the first two principal components of genotype matrix for SNPs.

Author Contributions: Conceptualization, Y.L., Y.Z. (Yuling Zhang), G.C. and E.Z.; methodology, Y.Z. (Yuling Zhang) and Y.L.; software, F.Z.; validation, Y.Z. (Yuexin Zhan) and Z.Y.; formal analysis, Z.F. and X.M.; investigation, Z.Y. and Y.L.; resources, E.Z.; data curation, G.C. and J.Y.; writing—original draft preparation, Y.L., Y.Z. (Yuexin Zhan), F.Z., Z.Y., Z.Z. and L.L.; writing—review and editing, Y.L., Y.Z. (Yuling Zhang), F.Z., E.Z., J.Y., Z.Z., L.L. and G.C.; visualization, Y.Z. (Yuexin Zhan), X.M. and Z.F.; supervision, Z.W., L.L. and Z.Z.; project administration, Z.W. and G.C.; funding acquisition, E.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Research and Development Project of Guangdong Province (Grant No. 2022B0202090002) and the Innovation Team Project of Modern Agricultural Industrial Technology System of Guangdong Province (2023KJ126).

Institutional Review Board Statement: All animals used in this study were treated in accordance with the guidelines for the use of laboratory animals of the Ministry of Agriculture of China and with the approval of South China Agricultural University (Guangzhou, China).

Informed Consent Statement: Not applicable.

Data Availability Statement: The genotype and phenotypic data analyzed in this study have not been made public at this time, as additional analytical studies will be conducted in the future. It can be obtained by contacting the corresponding author's email address if reasonably requested.

Acknowledgments: Thanks to Wens Foodstuff Group Co., Ltd. (Guangdong, China) and Guangdong Guangken Animal Husbandry Group Co., Ltd. for their assistance in sample collection.

Conflicts of Interest: The authors declare no conflict of interest. Zhenfang Wu and Gengyuan Cai were employed by company Guangdong Zhongxin Breeding Technology Co., Ltd. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Samorè, A.B.; Fontanesi, L. Genomic Selection in Pigs: State of the Art and Perspectives. *Ital. J. Anim. Sci.* **2016**, *15*, 211–232. <https://doi.org/10.1080/1828051X.2016.1172034>.
2. Meuwissen, T.H.E.; Hayes, B.J.; Goddard, M.E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* **2001**, *157*, 1819–1829. <https://doi.org/10.1093/genetics/157.4.1819>.
3. Goddard, M.E.; Hayes, B.J. Genomic Selection: Genomic Selection. *J. Anim. Breed. Genet.* **2007**, *124*, 323–330. <https://doi.org/10.1111/j.1439-0388.2007.00702.x>.




4. Pszczola, M.; Calus, M.P.L. Updating the Reference Population to Achieve Constant Genomic Prediction Reliability across Generations. *Animal* **2016**, *10*, 1018–1024. <https://doi.org/10.1017/S1751731115002785>.
5. VanRaden, P.M.; Van Tassell, C.P.; Wiggans, G.R.; Sonstegard, T.S.; Schnabel, R.D.; Taylor, J.F.; Schenkel, F.S. Invited Review: Reliability of Genomic Predictions for North American Holstein Bulls. *J. Dairy Sci.* **2009**, *92*, 16–24. <https://doi.org/10.3168/jds.2008-1514>.
6. Christensen, O.F.; Madsen, P.; Nielsen, B.; Ostensen, T.; Su, G. Single-Step Methods for Genomic Evaluation in Pigs. *Animal* **2012**, *6*, 1565–1571. <https://doi.org/10.1017/S1751731112000742>.
7. Sanchez, M.-P.; Govignon-Gion, A.; Croiseau, P.; Fritz, S.; Hozé, C.; Miranda, G.; Martin, P.; Barbat-Leterrier, A.; Letaïef, R.; Rocha, D.; et al. Within-Breed and Multi-Breed GWAS on Imputed Whole-Genome Sequence Variants Reveal Candidate Mutations Affecting Milk Protein Composition in Dairy Cattle. *Genet. Sel. Evol.* **2017**, *49*, 68. <https://doi.org/10.1186/s12711-017-0344-z>.
8. Song, H.; Ye, S.; Jiang, Y.; Zhang, Z.; Zhang, Q.; Ding, X. Using Imputation-Based Whole-Genome Sequencing Data to Improve the Accuracy of Genomic Prediction for Combined Populations in Pigs. *Genet. Sel. Evol.* **2019**, *51*, 58. <https://doi.org/10.1186/s12711-019-0500-8>.
9. Pérez-Enciso, M.; Rincón, J.C.; Legarra, A. Sequence- vs. Chip-Assisted Genomic Selection: Accurate Biological Information Is Advised. *Genet. Sel. Evol.* **2015**, *47*, 43. <https://doi.org/10.1186/s12711-015-0117-5>.
10. Larmer, S.G.; Sargolzaei, M.; Brito, L.F.; Ventura, R.V.; Schenkel, F.S. Novel Methods for Genotype Imputation to Whole-Genome Sequence and a Simple Linear Model to Predict Imputation Accuracy. *BMC Genet.* **2017**, *18*, 120. <https://doi.org/10.1186/s12863-017-0588-1>.
11. Heidaritabar, M.; Calus, M.P.L.; Megens, H.-J.; Vereijken, A.; Groenen, M.A.M.; Bastiaansen, J.W.M. Accuracy of Genomic Prediction Using Imputed Whole-Genome Sequence Data in White Layers. *J. Anim. Breed. Genet.* **2016**, *133*, 167–179. <https://doi.org/10.1111/jbg.12199>.
12. Ni, G.; Caverio, D.; Fangmann, A.; Erbe, M.; Simianer, H. Whole-Genome Sequence-Based Genomic Prediction in Laying Chickens with Different Genomic Relationship Matrices to Account for Genetic Architecture. *Genet. Sel. Evol.* **2017**, *49*, 8. <https://doi.org/10.1186/s12711-016-0277-y>.
13. Yan, G.; Qiao, R.; Zhang, F.; Xin, W.; Xiao, S.; Huang, T.; Zhang, Z.; Huang, L. Imputation-Based Whole-Genome Sequence Association Study Rediscovered the Missing QTL for Lumbar Number in Sutan Pigs. *Sci. Rep.* **2017**, *7*, 615. <https://doi.org/10.1038/s41598-017-00729-0>.
14. Druet, T.; Macleod, I.M.; Hayes, B.J. Toward Genomic Prediction from Whole-Genome Sequence Data: Impact of Sequencing Design on Genotype Imputation and Accuracy of Predictions. *Heredity* **2014**, *112*, 39–47. <https://doi.org/10.1038/hdy.2013.13>.
15. Van Binsbergen, R.; Calus, M.P.L.; Bink, M.C.A.M.; Van Eeuwijk, F.A.; Schrooten, C.; Veerkamp, R.F. Genomic Prediction Using Imputed Whole-Genome Sequence Data in Holstein Friesian Cattle. *Genet. Sel. Evol.* **2015**, *47*, 71. <https://doi.org/10.1186/s12711-015-0149-x>.
16. Ye, S.; Gao, N.; Zheng, R.; Chen, Z.; Teng, J.; Yuan, X.; Zhang, H.; Chen, Z.; Zhang, X.; Li, J.; et al. Strategies for Obtaining and Pruning Imputed Whole-Genome Sequence Data for Genomic Prediction. *Front. Genet.* **2019**, *10*, 673. <https://doi.org/10.3389/fgene.2019.00673>.
17. Brøndum, R.F. Quantitative Trait Loci Markers Derived from Whole Genome Sequence Data Increases the Reliability of Genomic Prediction. *J. Dairy Sci.* **2015**, *98*, 4107–4116.
18. Veerkamp, R.F.; Bouwman, A.C.; Schrooten, C.; Calus, M.P.L. Genomic Prediction Using Preselected DNA Variants from a GWAS with Whole-Genome Sequence Data in Holstein–Friesian Cattle. *Genet. Sel. Evol.* **2016**, *48*, 95. <https://doi.org/10.1186/s12711-016-0274-1>.
19. Zhang, Z.; Ober, U.; Erbe, M.; Zhang, H.; Gao, N.; He, J.; Li, J.; Simianer, H. Improving the Accuracy of Whole Genome Prediction for Complex Traits Using the Results of Genome Wide Association Studies. *PLoS ONE* **2014**, *9*, e93017. <https://doi.org/10.1371/journal.pone.0093017>.
20. Gebreyesus, G.; Bovenhuis, H.; Lund, M.S.; Poulsen, N.A.; Sun, D.; Buitenhuis, B. Reliability of Genomic Prediction for Milk Fatty Acid Composition by Using a Multi-Population Reference and Incorporating GWAS Results. *Genet. Sel. Evol.* **2019**, *51*, 16. <https://doi.org/10.1186/s12711-019-0460-z>.
21. Warburton, C.L.; Engle, B.N.; Ross, E.M.; Costilla, R.; Moore, S.S.; Corbet, N.J.; Allen, J.M.; Laing, A.R.; Fordyce, G.; Lyons, R.E.; et al. Use of Whole-Genome Sequence Data and Novel Genomic Selection Strategies to Improve Selection for Age at Puberty in Tropically-Adapted Beef Heifers. *Genet. Sel. Evol.* **2020**, *52*, 28. <https://doi.org/10.1186/s12711-020-00547-5>.
22. Edwards, S.M.; Sørensen, I.F.; Sarup, P.; Mackay, T.F.C.; Sørensen, P. Genomic Prediction for Quantitative Traits Is Improved by Mapping Variants to Gene Ontology Categories in *Drosophila melanogaster*. *Genetics* **2016**, *203*, 1871–1883. <https://doi.org/10.1534/genetics.116.187161>.
23. Zhang, H.; Zhuang, Z.; Yang, M.; Ding, R.; Quan, J.; Zhou, S.; Gu, T.; Xu, Z.; Zheng, E.; Cai, G.; et al. Genome-Wide Detection of Genetic Loci and Candidate Genes for Body Conformation Traits in Duroc × Landrace × Yorkshire Crossbred Pigs. *Front. Genet.* **2021**, *12*, 664343. <https://doi.org/10.3389/fgene.2021.664343>.
24. Misztal, I.; Lourenco, D.; Aguilar, I.; Legarra, A.; Vitezica, Z. *Manual for BLUPF90 Family of Programs*; University of Georgia: Athens, GA, USA, 2014.
25. Wang, Y.; Ding, X.; Tan, Z.; Ning, C.; Xing, K.; Yang, T.; Pan, Y.; Sun, D.; Wang, C. Genome-Wide Association Study of Piglet Uniformity and Farrowing Interval. *Front. Genet.* **2017**, *8*, 194. <https://doi.org/10.3389/fgene.2017.00194>.

26. Chang, C.C.; Chow, C.C.; Tellier, L.C.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets. *Gigascience* **2015**, *4*, 7. <https://doi.org/10.1186/s13742-015-0047-8>.
27. Ding, R.; Savegnago, R.; Liu, J.; Long, N.; Tan, C.; Cai, G.; Zhuang, Z.; Wu, J.; Yang, M.; Qiu, Y.; et al. The SWine IMputation (SWIM) Haplotype Reference Panel Enables Nucleotide Resolution Genetic Mapping in Pigs. *Commun. Biol.* **2023**, *6*, 577. <https://doi.org/10.1038/s42003-023-04933-9>.
28. Yang, J.; Benyamin, B.; McEvoy, B.P.; Gordon, S.; Henders, A.K.; Nyholt, D.R.; Madden, P.A.; Heath, A.C.; Martin, N.G.; Montgomery, G.W.; et al. Common SNPs Explain a Large Proportion of the Heritability for Human Height. *Nat. Genet.* **2010**, *42*, 565–569. <https://doi.org/10.1038/ng.608>.
29. Yang, J.; Lee, S.H.; Goddard, M.E.; Visscher, P.M. GCTA: A Tool for Genome-Wide Complex Trait Analysis. *Am. J. Hum. Genet.* **2011**, *88*, 76–82. <https://doi.org/10.1016/j.ajhg.2010.11.011>.
30. Yin, L.; Zhang, H.; Tang, Z.; Yin, D.; Fu, Y.; Yuan, X.; Li, X.; Liu, X.; Zhao, S. HIBLUP: An Integration of Statistical Models on the BLUP Framework for Efficient Genetic Evaluation Using Big Genomic Data. *Nucleic Acids Res.* **2023**, *51*, 3501–3512. <https://doi.org/10.1093/nar/gkad074>.
31. Su, G.; Christensen, O.F.; Janss, L.; Lund, M.S. Comparison of Genomic Predictions Using Genomic Relationship Matrices Built with Different Weighting Factors to Account for Locus-Specific Variances. *J. Dairy Sci.* **2014**, *97*, 6547–6559. <https://doi.org/10.3168/jds.2014-8210>.
32. De Las Heras-Saldana, S.; Lopez, B.I.; Moghaddar, N.; Park, W.; Park, J.; Chung, K.Y.; Lim, D.; Lee, S.H.; Shin, D.; Van Der Werf, J.H.J. Use of Gene Expression and Whole-Genome Sequence Information to Improve the Accuracy of Genomic Prediction for Carcass Traits in Hanwoo Cattle. *Genet. Sel. Evol.* **2020**, *52*, 54. <https://doi.org/10.1186/s12711-020-00574-2>.
33. Ye, S.; Li, J.; Zhang, Z. Multi-Omics-Data-Assisted Genomic Feature Markers Preselection Improves the Accuracy of Genomic Prediction. *J. Anim. Sci. Biotechnol.* **2020**, *11*, 109. <https://doi.org/10.1186/s40104-020-00515-5>.
34. MacLeod, I.M.; Bowman, P.J.; Vander Jagt, C.J.; Haile-Mariam, M.; Kemper, K.E.; Chamberlain, A.J.; Schrooten, C.; Hayes, B.J.; Goddard, M.E. Exploiting Biological Priors and Sequence Variants Enhances QTL Discovery and Genomic Prediction of Complex Traits. *BMC Genom.* **2016**, *17*, 144. <https://doi.org/10.1186/s12864-016-2443-6>.
35. Moghaddar, N.; Khansefid, M.; van der Werf, J.H.J.; Bolormaa, S.; Duijvesteijn, N.; Clark, S.A.; Swan, A.A.; Daetwyler, H.D.; MacLeod, I.M. Genomic Prediction Based on Selected Variants from Imputed Whole-Genome Sequence Data in Australian Sheep Populations. *Genet. Sel. Evol.* **2019**, *51*, 72. <https://doi.org/10.1186/s12711-019-0514-2>.
36. Lopez, B.I.M.; An, N.; Srikanth, K.; Lee, S.; Oh, J.-D.; Shin, D.-H.; Park, W.; Chai, H.-H.; Park, J.-E.; Lim, D. Genomic Prediction Based on SNP Functional Annotation Using Imputed Whole-Genome Sequence Data in Korean Hanwoo Cattle. *Front. Genet.* **2021**, *11*, 603822. <https://doi.org/10.3389/fgene.2020.603822>.
37. Song, H.; Hu, H. Strategies to Improve the Accuracy and Reduce Costs of Genomic Prediction in Aquaculture Species. *Evol. Appl.* **2022**, *15*, 578–590. <https://doi.org/10.1111/eva.13262>.
38. Zhang, C.; Kemp, R.A.; Stothard, P.; Wang, Z.; Boddicker, N.; Krivushin, K.; Dekkers, J.; Plastow, G. Genomic Evaluation of Feed Efficiency Component Traits in Duroc Pigs Using 80K, 650K and Whole-Genome Sequence Variants. *Genet. Sel. Evol.* **2018**, *50*, 14. <https://doi.org/10.1186/s12711-018-0387-9>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Identification of Candidate Genes for Economically Important Carcass Cutting in Commercial Pigs through GWAS

Fuchen Zhou ^{1,†} , Jianping Quan ^{1,†}, Donglin Ruan ¹, Yibin Qiu ¹, Rongrong Ding ¹ , Cineng Xu ¹, Yong Ye ¹, Gengyuan Cai ^{1,2,3}, Langqing Liu ¹, Zebin Zhang ¹, Jie Yang ^{1,2} , Zhenfang Wu ^{1,2,3,4,*} and Enqin Zheng ^{1,2,*}

¹ College of Animal Science and National Engineering Research Center for Breeding Swine Industry, South China Agricultural University, Guangzhou 510642, China; zfc17854225519@163.com (F.Z.); qjp_scau@outlook.com (J.Q.); ruandl@stu.scau.edu.cn (D.R.); 13422157044qyb@gmail.com (Y.Q.); drr_scau@foxmail.com (R.D.); cnxu@stu.scau.edu.cn (C.X.); yinhun0517@163.com (Y.Y.); cgy0415@163.com (G.C.); langqing.liu@scau.edu.cn (L.L.); zbzhang@scau.edu.cn (Z.Z.); jieyang2012@hotmail.com (J.Y.)

² Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, South China Agricultural University, Guangzhou 510642, China

³ Guangdong Zhongxin Breeding Technology Co., Ltd., Guangzhou 510642, China

⁴ Yunfu Subcenter of Guangdong Laboratory for Lingnan Modern Agriculture, Yunfu 527400, China

* Correspondence: wzfemail@163.com (Z.W.); eqzheng@scau.edu.cn (E.Z.)

[†] These authors contributed equally to this work.

Simple Summary: Tenderloin and rib weight are important components of the economic value of pig carcasses, and selling them separately after fine segmentation further enhances the economic value of the carcasses. This study represents one of the rare attempts to conduct a genome-wide analysis focused on the economic value of pig carcasses, utilizing post-slaughter carcass phenotype values and genotype data to identify genetic variation regions. Through our investigation, we have identified several promising candidate regions and genes that have the potential to contribute valuable insights for breeding strategies and marker-assisted selection in pig production.



Citation: Zhou, F.; Quan, J.; Ruan, D.; Qiu, Y.; Ding, R.; Xu, C.; Ye, Y.; Cai, G.; Liu, L.; Zhang, Z.; et al. Identification of Candidate Genes for Economically Important Carcass Cutting in Commercial Pigs through GWAS. *Animals* **2023**, *13*, 3243. <https://doi.org/10.3390/ani13203243>

Academic Editor: Martino Cassandro

Received: 12 September 2023

Revised: 8 October 2023

Accepted: 16 October 2023

Published: 18 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: During the process of pork production, the carcasses of pigs are divided and sold, which provides better economic benefits and market competitiveness for pork production than selling the carcass as a whole. Due to the significant cost of post-slaughter phenotypic measurement, the genetic architecture of tenderloin weight (TLNW) and rib weight (RIBW)—important components of pig carcass economic value—remain unknown. In this study, we conducted genome-wide association studies (GWAS) for TLNW and RIBW traits in a population of 431 Duroc × Landrace × Yorkshire (DLY) pigs. In our study, the most significant single nucleotide polymorphism (SNP) associated with TLNW was identified as ASGA0085853 (3.28 Mb) on Sus scrofa chromosome 12 (SSC12), while for RIBW, it was Affx-1115046258 (172.45 Mb) on SSC13. Through haplotype block analysis, we discovered a novel quantitative trait locus (QTL) associated with TLNW, spanning a 5 kb region on SSC12, and a novel RIBW-associated QTL spanning 1.42 Mb on SSC13. Furthermore, we hypothesized that three candidate genes, *TIMP2* and *EML1*, and *SMN1*, are associated with TLNW and RIBW, respectively. Our research not only addresses the knowledge gap regarding TLNW, but also serves as a valuable reference for studying RIBW. The identified SNP loci strongly associated with TLNW and RIBW may prove useful for marker-assisted selection in pig breeding programs.

Keywords: pigs; fine segmentation; tenderloin weight; rib weight; genome-wide association studies

1. Introduction

Pork is the most widely consumed meat globally and serves as a crucial protein source in human nutrition [1,2]. Due to diverse dietary preferences, pork consumption varies across regions, resulting in substantial price discrepancies for different pork cuts.

Carcass segmentation enables the separate sale of distinct parts, thereby enhancing the overall market value of the carcass and providing better economic benefits and market competitiveness for pork production [3,4]. In the Chinese market, tenderloin represents the highest-priced fresh meat [5,6], and customers particularly appreciate tenderloin for its attributes such as juiciness, low fat content, and high protein content [7,8]. Ribs are important components of the pig carcass and provide a maximum reflection of the pig's economic value. Therefore, tenderloin weight (TLNW) and rib weight (RIBW) as the quantitative traits are the key indicators of pig carcass segmentation.

However, the high costs associated with slaughter testing and the challenges in data collection have posed significant obstacles in unraveling these genetic mechanisms. Limited research has been conducted specifically focusing on post-slaughter carcass traits. Rib number, as an indirect indicator of rib weight, is influenced by a diverse array of genetic factors. Notably, the *VRTN* gene, associated with vertebral development, has undergone extensive investigation and emerged as a promising candidate gene for regulating rib number. It is localized within the quantitative trait loci (QTL) region of swine chromosome 7 [6,9–13]. Additionally, previous studies have shed light on other factors influencing rib growth, including the *LTBP2* gene, which has demonstrated its capability to enhance rib number in knockout mice [5]. In their study on the regulation of muscle development, Van Laere et al. [14] discovered that the *IGF2* gene plays a crucial role in muscle development in pigs. Oczkiewicz et al. [15] found that the *IGF2* gene leads to a significant increase in tenderloin weight (11 ± 0.01 g). Furthermore, Burgos et al. [16] reported the capacity of the *IGF2* gene to enhance tenderloin muscle tissue in pigs. Xie et al. [13] identified the *CD96* gene as a potential candidate gene for influencing tenderloin. However, there is currently limited genetic structural analysis directly targeting post-slaughter carcass traits, and there are still many uncertainties in the genetic mechanisms underlying these traits.

In recent years, the rapid advancement and application of high-throughput sequencing technology have propelled genome-wide association studies (GWAS) as a robust strategy for identifying genetic variations associated with complex traits. This approach has found extensive use in the fields of husbandry [17–19] and human disease research [20,21]. Compared to previous studies employing microsatellite molecular markers [22,23], the combination of GWAS with single nucleotide polymorphism (SNP) molecular marker technology offers greater accuracy in detecting QTL intervals [24]. GWAS has also been extensively employed to investigate the genetic variation and diversity underlying economically important traits in pigs [25,26]. In previous studies [27,28], GWAS, utilizing genotype information from the GeneSeek Porcine 50K SNP Chip, has successfully identified numerous significant QTLs and candidate genes associated with important economic traits.

To further pinpoint the key loci influencing carcass segmentation traits, we performed a GWAS on the post-slaughter traits of TLNW and RIBW in a cohort of 431 Duroc \times Landrace \times Yorkshire (DLY) pigs. The aim of this study was to unravel the genetic architecture underlying carcass segmentation traits and facilitate the rapid development of molecular breeding in pigs.

2. Materials and Methods

2.1. Ethics Statement

The animals and experimental procedures used in this study were handled following the guidelines set forth by the Animal Care and Use Committee of the South China Agricultural University (SCAU) (Guangzhou, China). The ethics committee of SCAU approved all animal experiments. The experimental animals were not anesthetized or euthanized during this study.

2.2. Samples and Phenotype Data

In the present study, we collected 431 three-way crossbred DLY pigs from Wens Foodstuff Group Co., Ltd. (Yunfu, China). All pigs were subjected to the same growth and feeding conditions. After that, the unified slaughtering was carried out according to the standard slaughtering flow at (110 ± 5) kg body mass. Pigs fasted for 24 h before slaughter,

were provided with free drinking water, and were subjected to neither beating nor driving before slaughtering, so as not to affect the meat slaughtering experiment carried out in the slaughterhouse. All pigs were divided into five batches and slaughtered in a commercial abattoir in Chifeng, Inner Mongolia. Electric shock anesthesia and heart bloodletting were used for slaughtering, and fine segmentation of each pork carcass was subsequently carried out. After fine segmentation, the weight of either the tenderloin or the ribs was recorded, as TLNW and RIBW, respectively (431 pigs for TLNW, 408 pigs for RIBW). Further details on the segmentation position are shown in Figure 1.

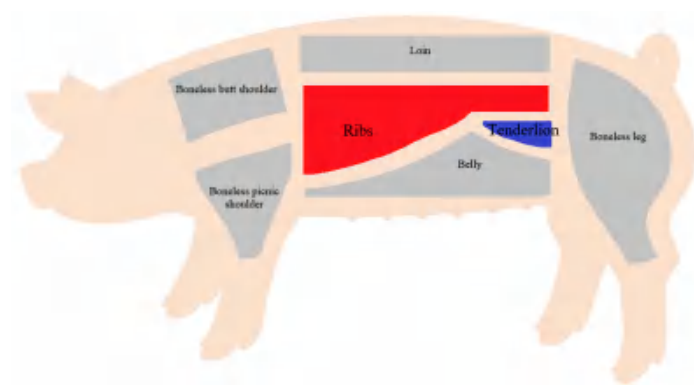


Figure 1. Schematic diagram of carcass cutting after pig slaughter.

2.3. SNP Genotyping and Quality Control

The genomic DNA needed in this experiment was isolated and extracted from the ear tissue of 431 pigs using the standard phenol/chloroform method. All 431 DNA samples were subjected to DNA quality control according to light absorption ratio (A260/280 and A260/230), gel electrophoresis, and DNA concentration of 50 ng/μL. The GeneSeek pig 50 K SNP chip was used for genotyping with a total of 50,643 SNPs. The genotype quality control of the 431 DLY pigs was conducted using PLINK v1.9 software [29]. Individuals with a call rate of less than 95% and SNPs with a call rate of less than 90% and a minor allele frequency of less than 0.01 were deleted. SNPs that failed the Hardy–Weinberg equilibrium test ($p < 10^{-6}$) and were unmapped or located on the sex chromosome were also removed. After QC, 4188 SNPs not located on autosome chromosomes were discarded. Moreover, we removed 211 (TLNW) and 213 (RIBW) SNPs because of missing genotype data; 11,202 SNPs due to failing the Hardy–Weinberg exact test; and 64 SNPs due to their minor allele threshold. All animals passed the QC (431 for TLNW and 408 for RIBW). Finally, 34,978 TLNW SNPs and 34,976 RIBW SNPs were retained for subsequent analysis.

2.4. Population Structure and Single-Locus GWAS Analysis

Population stratification is one of the main reasons for unreliable GWAS results, as it can cause false positive results. Principal component analysis (PCA) and LD analysis were performed using the SNPs that met the QC standards to investigate the population structure. PCA was performed with GCTA software (version 1.93.2 beta) [30]. In addition, the quantile–quantile (Q–Q) plot and inflation factor (λ) were obtained using the qqman package in R software (version 4.1.2).

GEMMA software (version 0.98.5) was used to implement a Mixed Linear Model (MLM) for single-locus GWAS of TLNW and RIBW [31]. GEMMA calculated the genome correlation matrix (GRM) between individuals in each population to illustrate the population structure. The first five principal components calculated by the GCTA tool are embedded into the correlation analysis model as covariables to eliminate the mixed influence of population structure [32]. The model for testing the allelic effects of TLNW and RIBW invoked by each SNP to GEMMA is as follows:

$$y = W\alpha + X\beta + u + \varepsilon$$

where y represents a vector of TLNW and RIBW; W is the incidence matrix of covariates, including fixed effects of the top three eigenvectors of sex, live weight, slaughter batch, and the top five principal component from PCA analysis; α represents the vector of corresponding coefficients including the intercept; X is the vector of all marker genotypes; β specifies the corresponding effect size of the marker size; u is the vector of random effects, with $u \sim \text{MVN}_n(0, \lambda\tau^{-1}K)$; ε is the vector of random residuals, with $\varepsilon \sim \text{MVN}_n(0, \tau^{-1}In)$; λ signifies the ratio between two variance components; τ^{-1} is the variance of the residual errors; K is the GRM; I is an $n \times n$ identity matrix; and n refers to the number of pigs. In the study, Bonferroni correction was used to determine the threshold p values of single-locus GWAS. At a stringent genome-wide Bonferroni threshold, $p < (0.05/N)$. At a more lenient threshold, $p < (1/N)$ for chromosome-wide (suggestive) associations, and N means the number of SNPs [25]. Haploview v4.2 software was used to perform haplotype block analysis to estimate the LD pattern of significant SNPs in an LD block [33]. LD among SNPs were estimated as the squared correlation (r^2) of alleles with a window size of 1000 kb.

The model in GCTA software, as following, is used to estimate the SNP-based heritability and the phenotypic variance explained by genome-wide SNPs (based on SNP inheritance), the proportion of phenotypic variation explained by significant SNPs:

$$y = X\beta + g + \varepsilon \text{ with } \text{var}(y) = Ag\sigma_g^2 + I\sigma_\varepsilon^2$$

where y is the vector of tenderloin weight or ribs weight; β is the vector including fixed effects; X is an incidence matrix for β ; g is the vector of the aggregate effects of all the qualified 50K SNPs for the pigs within one population; I is the identity matrix; Ag is the genomic relatedness matrix estimated by these SNPs; σ_g^2 is the additive genetic variance captured by either the genome-wide SNPs or the selected SNPs; and σ_ε^2 is the residual variance. The heritability and the phenotypic variance explained by the SNPs can be estimated using the model simply described as $h^2 = \sigma_g^2 / \sigma_p^2$, where σ_p^2 (total phenotypic variance) is the sum of σ_g^2 and σ_ε^2 .

2.5. Identification of Candidate Genes and Functional Analysis

All SNPs refer to the latest version of the *Sus scrofa* 11.1 genome (http://ensembl.org/Sus_scrofa/Info/Index, accessed on 3 August 2023). Functional gene annotation (v105) was downloaded in GFF3 format from the Ensembl website (http://ftp.ensembl.org/pub/release-105/gff3/sus_scrofa/, accessed on 3 August 2023). The R package BioMart (version 2.56.1) [34] efficiently retrieved functional genes. KEGG and GO analyses were conducted using KOBAS 3.0 [35] to investigate the functions of all candidate genes. Enriched terms with a significance threshold of p value < 0.05 were selected to further explore the genes invoked in pathway and biological processes. Subsequently, we employed REVIGO (<http://revigo.irb.hr>, accessed on 8 August 2023) in conjunction with the *Mus musculus* database to eliminate GO term redundancy (medium threshold, 0.7) and cluster the remaining terms in a 2D space [36,37]. This space was derived by applying multidimensional scaling to a matrix of GO terms with semantic similarities. The Mouse Genome Informatics website (<https://www.informatics.jax.org/>, accessed on 10 August 2023), GeneCards (<http://www.genecards.org/>, accessed on 10 August 2023), and Ensembl (www.ensembl.org/biomart/martview, accessed on 10 August 2023) were used to query gene functions.

3. Results

3.1. Phenotypic Variation and SNP Genotyping

Table 1 presents various phenotypic and genetic parameters for TLNW and RIBW traits, including animal count, mean, standard deviation, maximum, minimum, coefficient of variation (CV), and heritability. On average, TLNW and RIBW were 0.46 ± 0.08 kg and 4.51 ± 0.56 kg, respectively. The CV values for TLNW and RIBW were 17.40% and 12.42%, respectively. The distribution and visualization of the SNP dataset across chromosomes are summarized in Figure S1, Tables S1 and S2. These SNPs were roughly proportionally

distributed across all 18 chromosomes of pigs, with the longest chromosome having the highest number of SNPs. Importantly, the SNP-based heritability (including standard errors) for TLNW and RIBW were 0.42 (0.11) and 0.22 (0.09), respectively.

Table 1. Summary statistics of tenderloin weight and ribs weight.

Trait	N ³	Mean (\pm SD)/kg ⁴	Min/kg ⁵	Max/kg ⁶	C.V./% ⁷	h ² (\pm SE) ⁸
TLNW ¹	431	0.46 \pm 0.08	0.24	0.69	17.40	0.42 \pm 0.11
RIBW ²	408	4.51 \pm 0.56	2.87	6.17	12.42	0.22 \pm 0.09

¹ Tenderloin weight (TLNW). ² Ribs weight (RIB). ³ Number of animals (N). ⁴ Standard deviations (SD). ⁵ Minimum (Min). ⁶ Maximum (Max). ⁷ Coefficient of variation (C.V.). ⁸ Heritability (standard error) value (h² (\pm SE)).

3.2. Single-Locus GWAS for TLNW and RIBW

Population stratification is a significant factor contributing to the unreliability of GWAS data, as it can result in false positive findings. The quantile–quantile (Q–Q) plots serve as a valuable tool for assessing the presence of population stratification [38]. In our study, the genomic inflation factors (λ) for TLNW and RIBW GWAS were determined to be 1.04 and 1.00, respectively (Figure 2). These values suggest that the TLNW and RIBW data obtained from the DLY population in our study are not influenced by population stratification.

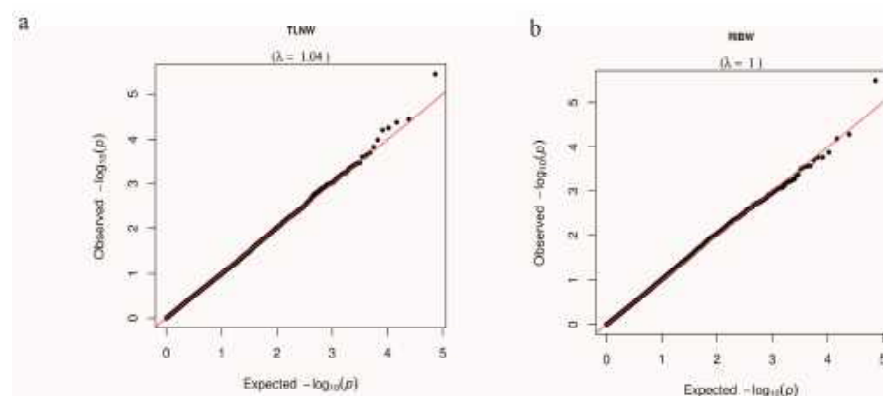


Figure 2. The Q–Q plots of TLNW (a) and RIBW (b) for DLY population. The Q–Q plot shows the observed versus expected $-\log_{10} p$ value. The red line represents observed values equal to expected values.

The mixed model was used to perform a single marker test, aiming to identify genetic markers associated with the TLNW and RIBW traits. The significant SNPs distinguished by single-locus GWAS for TLNW and RIBW are shown in Figure 3 and Table 2. The chromosome-wide (suggestive) Bonferroni-corrected thresholds of TLNW and RIBW were $p < 2.86 \times 10^{-5}$ (1/34,978) and $p < 2.86 \times 10^{-5}$ (1/34,976), respectively. Furthermore, two suggestive SNPs (ASGA0085853 and ALGA0112188) were found to be associated with TLNW on *Sus scrofa* chromosome (SSC) 12 and SSC7, respectively, while one suggestive SNP (Affx-1115046258) was related to RIBW on SSC13. Moreover, Figure 3a shows that the most significant cluster is on SSC12, indicating a strong signal. The most significant SNPs for TLNW and for RIBW were ASGA0085853 and Affx-1115046258, respectively. Additionally, on SSC12, ASGA0085853 is positioned at 3.28 Mb with a minor allele frequency (MAF) of 0.306, yielding a $-\log_{10}(p\text{-value})$ of 5.16. On SSC13, Affx-1115046258 is located at 172.45 Mb with a MAF of 0.268 and a $-\log_{10}(p\text{-value})$ of 5.31. The most significant SNPs (ASGA0085853 and Affx-1115046258) for the above characterized haplotype block explained 4.88% and 5.19% of the phenotypic variance for TLNW and RIBW, respectively. Among these, carriers of the A allele (AA and AG) of ASGA0085853 had significantly greater loin weight than those with the GG genotype, with highly significant phenotypic differences observed among the three genotypes (Figure 4a). However, for Affx-1115046258, there were no significant differences in phenotype among the three genotypes (Figure S2).

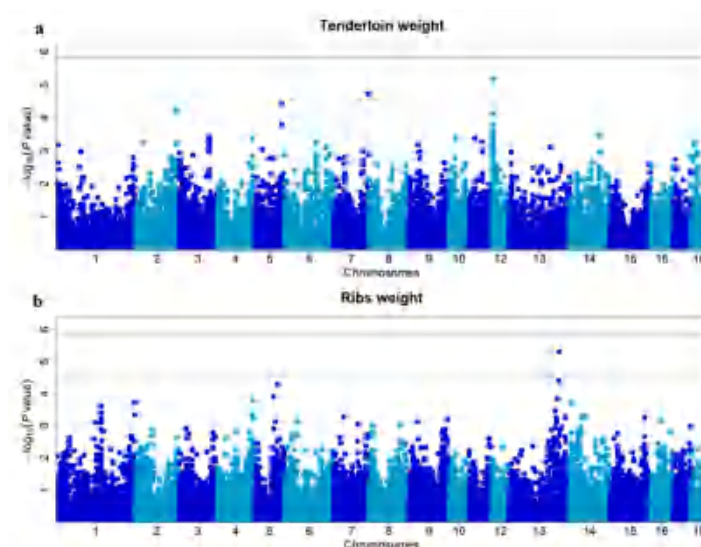


Figure 3. Manhattan plots of GWAS for TLNW (a) and RIBW (b) in the DLY population. In the Manhattan plots, the solid and dashed lines represent the 5% genome-wide and chromosome-wide (suggestive) Bonferroni-corrected thresholds, respectively. The x-axis represents the chromosomes, and the y-axis represents the $-\log_{10}(p \text{ value})$.

Table 2. Significant SNPs for TLNW and RIBW in DLY Pigs.

Trait	SNP	SSC ¹	Position (bp)	EPV ²	MAF	<i>p</i> -Value	Distance ³	Nearest Gene
TLNW	ASGA0085853	12	3,284,259	4.88%	0.306	6.88×10^{-6}	within	<i>TIMP2</i>
	ALGA0112188	7	120,821,692	3.90%	0.325	1.92×10^{-5}	within	<i>EML1</i>
RIBW	Affx-115046258	13	172,454,121	5.19%	0.268	4.88×10^{-6}	150.8 kb	<i>ENSSSCG00000029127</i>

¹ Sus scrofa chromosome (SSC). ² Explained phenotypic variance (EPV). ³ The SNP located upstream/downstream of the nearest gene (Distance).

3.3. Effects of the QTL for TLNW and RIBW

Haploview v4.2 [33] can visualize the linkage disequilibrium (LD) and/or linkage between significant SNPs on the same chromosome, forming block and linkage value. The QTL regions recognized by Haploview v4.2 are shown in Figure 4. The leading SNPs (ASGA0085853, Affx-115046258) were mapped to two QTL regions spanning 5 kb and 1.42 Mb, respectively. For TLNW, one QTL region was identified on SSC12, which was composed of only two SNPs located between 3,284,259 and 3,289,920 bp (Figure 4b). In addition, there is very strong linkage between ASGA0085853 ($p\text{-value} = 6.88 \times 10^{-6}$) and ASGA0084858 ($p\text{-value} = 7.45 \times 10^{-5}$). For RIBW, the QTL region on SSC13 was composed of six SNPs; the most significant, Affx-115046258 (172.45 Mb), was linked closely with the other five SNPs (MARC0016316, WU_10.2_13_181846347, MARC0067784, CASI0008207, ALGA0072835) in the QTL region (Figure 4c). As illustrated in Figures 3b and 4c, CASI0008207 is the second most significant SNP in the GWAS results of the RIBW trait.

3.4. Candidate Genes Search and Functional Annotation

In the analysis of TLNW, within a range of 500 kb upstream and downstream of the significant SNPs, we annotated 11 and 13 protein-coding genes on SSC7 and SSC12, respectively (Table S3). Notably, both of the significant SNPs were located within the *TIMP2* and *EML1* genes. Our pathway enrichment analysis revealed several significantly enriched terms from the Kyoto Encyclopedia of Genes and Genomes (KEGG) and the Gene Ontology (GO) knowledgebase that are relevant to TLNW. These enriched terms include cellular division and protein translation (Figure 5a,b, Table S4). After conducting non-redundant GO analysis on all GO terms that exceeded the threshold using the REVIGO website, a

total of 39 GO terms were clustered (Figure 6a,b). The most prominent signaling pathway among them is negative regulation of ruffle assembly (GO:1900028). Subsequently, we employed the GeneCards—Mouse Genome Informatics databases—and conducted an extensive literature review to explore the functional roles of the identified genes. As a result, we identified a total of four candidate genes with potential relevance to TLNW. These genes, namely *YY1*, *EML1*, *CANT1*, and *TIMP2*, exhibit promising associations with TLNW based on their known functions and previous research findings. Furthermore, we identified the *SMN1* gene as a strong candidate gene for the RIBW trait. The *SMN1* gene is homologous to the *ENSSSCG00000029127* gene and is located downstream of the leading SNP by approximately 150.8 kb.

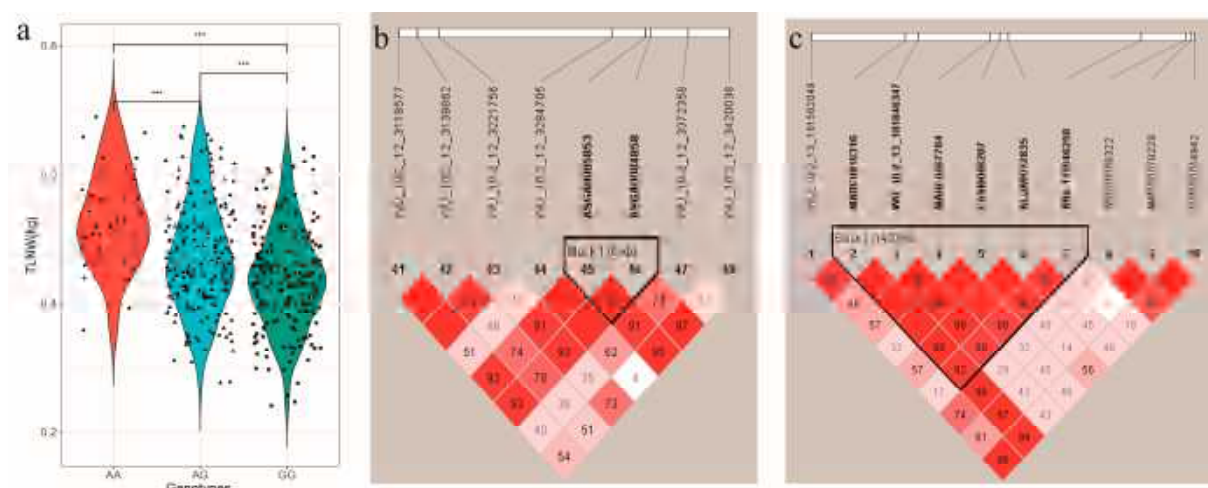


Figure 4. The plot (a) indicates the genotype effect plot of top SNP (ASGA008583) related to TLNW in 431 DLY pigs (** $p < 0.01$). Haplotype block for (b) TLNW and (c) RIBW in DLY pigs, respectively. Haplotype blocks are marked with triangles. Values in boxes are the linkage disequilibrium (r^2) between the SNP pairs. The haplotype blocks are colored in accordance with the standard Haploview color scheme: $LOD > 2$ and $D' = 1$, red; $LOD < 2$ and $D' < 1$, white (LOD is the log of the likelihood odds ratio, a measure of confidence in the value of D').

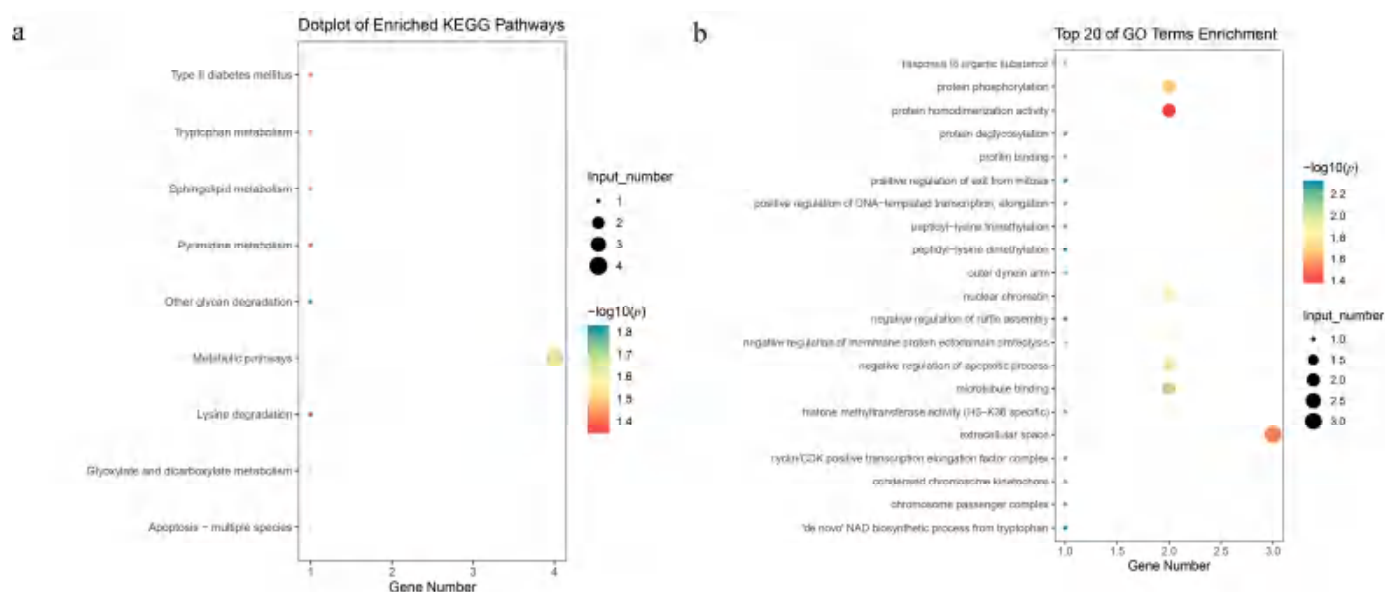
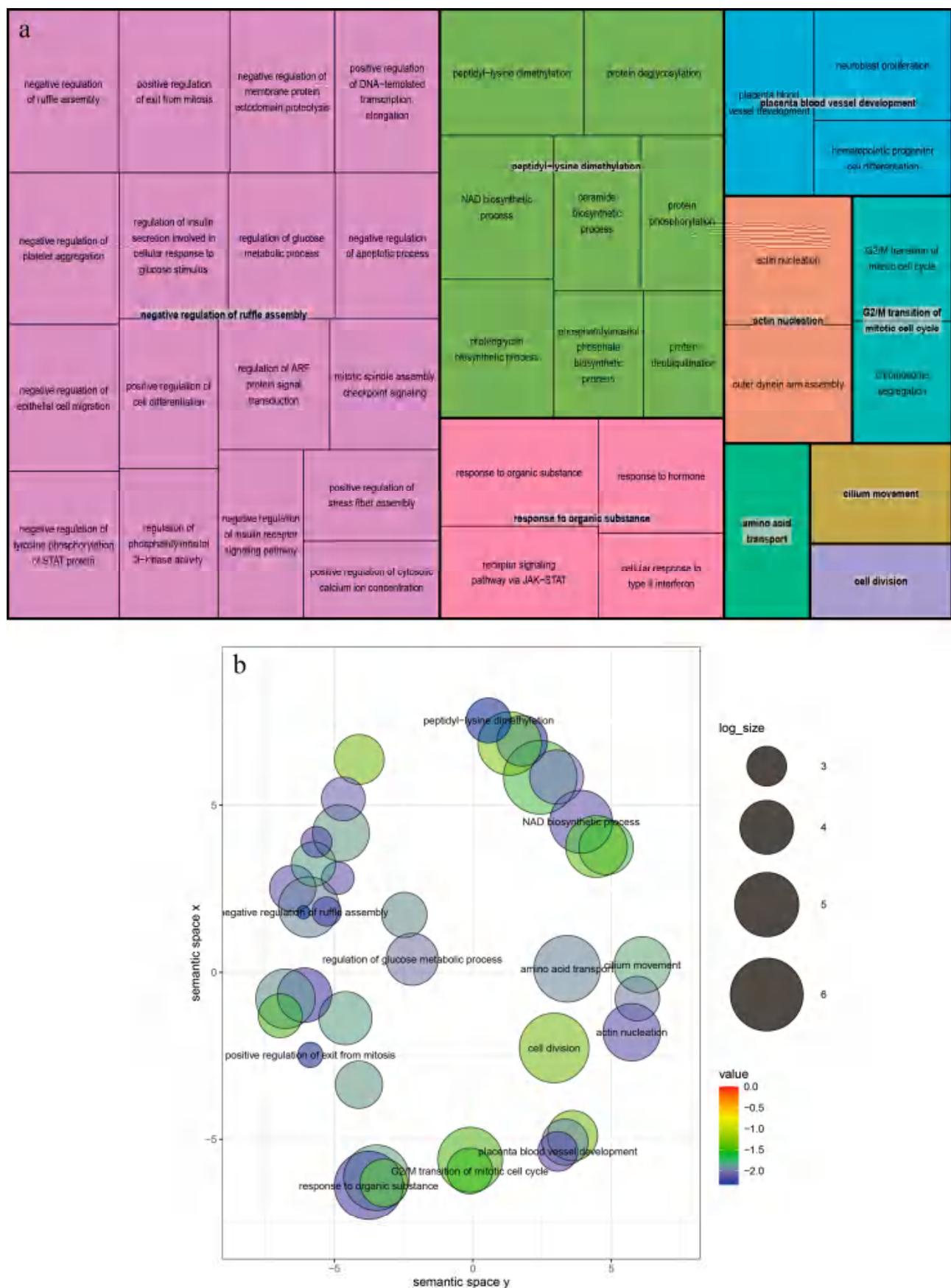


Figure 5. Significant KEGG pathways and GO terms associated with TLNW ($p < 0.05$). The plot (a) represents the KEGG pathway of the biological process for protein-coding genes within a 1 Mb region centered on the significant SNPs. The plot (b) shows the top 20 terms of the GO enrichment.



Each rectangle represents a representative cluster. These representatives are combined into “super-clusters,” representing loosely related terms and visualized using different colors. The size of the rectangles is adjusted to reflect the p value and frequency of the GO term in the *Mus musculus* GOA database. The plot (b) displays the scatter plot of representative clusters. The log size indicates the frequency of the GO term in the *Mus musculus* GOA database, with larger sizes indicating more common terms. The numerical value represents the $-\log_{10}(p \text{ value})$, with colors ranging from red to blue indicating increasing significance.

4. Discussion

4.1. Fine Segmentation and Sale of Pig Carcasses

According to the 2021 edition of the OECD-FAO Agricultural Outlook, the global meat supply is projected to expand, reaching 374 million tons by 2030 (https://www.oecd-ilibrary.org/agriculture-and-food/oecd-fao-agriculturaloutlook_19991142, accessed on 1 August 2022). While the demand for pork is gradually rising worldwide, many nations have preferences for specific varieties of meat from pig carcasses [39]. For example, Germans consume fresh shoulder pork more frequently than any other portion [40], while Denmark is known for its high frequency consumption of liver [40]. In South Korea, fresh pork belly is an extremely popular meat, accounting for 59% of the per capita consumption of approximately 100 g of meat per day [41]. Additionally, TLN and RIB are highly favored by Chinese people due to their excellent meat quality, despite being relatively expensive compared to other parts [42]. Slaughtering and selling pig carcasses through fine segmentation can satisfy different consumers' preferences for pork, showcase the true value of hogs, and fully explore the subsequent processes of existing animal husbandry production. Pig producers can enhance pig performance and increase economic value by utilizing genetic structure analysis of the TLNW and RIBW traits.

4.2. Genetic Loci and Candidate Genes for the TLNW Trait

The leading SNP (ASGA0085853) was annotated within the first intron of the *Metalloproteinase inhibitor 2 (TIMP2)* gene on SSC12 through *Sus scrofa* 11.1 delivered from the Ensembl database. In mice, deficiency of *TIMP2* leads to increased cardiac hypertrophy and subsequent heart enlargement [43]. In cows, mRNA expression of *TIMP2* is associated with intramuscular fat content and explains 32% of the variation in intramuscular fat [44]. *TIMP2* also affects feed conversion efficiency in cattle by regulating cell growth and proliferation networks [45,46]. Based on these findings, *TIMP2* may regulate the size of cardiomyocytes by influencing the molecular pathways of cell growth and proliferation, ultimately affecting the development of the TLNW trait. On SSC12, the candidate gene *CANT1* is located 79.8 kb upstream of the leading SNP and is associated with abnormal skeletal morphology and body size in mice [47]. Additionally, *EML1*—found on SSC7, a candidate gene associated with TLNW—is located within ALGA0112188. *EML1* is associated with brain overgrowth syndrome [48] and plays a crucial role in proper retinal lamination during cellular proliferation and development [49] in humans. The potential function of *EML1* suggested by this study requires further investigation for functional validation. On SSC7, the candidate gene, *YY1*, has been shown to be involved in cell proliferation and body size in mice [50]. It is suggested that *YY1* might regulate tenderloin growth through processes related to the proliferation and development of muscle cells. These results indicate that the *TIMP2* and *EML1* genes may play an important role in TLNW and should be considered strong candidate genes for this trait.

4.3. Genetic Loci and Candidate Genes for the RIBW Trait

After conducting haplotype block analysis, we identified *ENSSSCG00000049210* and *ENSSSCG00000050907* as noncoding genes within the QTL on SSC13. However, we discovered a protein-coding gene, *ENSSSCG00000029127*, located 150.8 kb downstream of the leading SNP (Affx-115046258). Interestingly, *ENSSSCG00000029127* exhibits homology to the protein encoded by the *SMN1* gene. Our literature search revealed that the

SMN1 gene is the determinant gene for spinal muscular atrophy in humans [51], a rare hereditary neuromuscular disease caused by deletion and/or mutation of *SMN1* [52]. The *SMN1* gene has been demonstrated to be associated with physiological conditions such as abnormal muscle physiology, decreased body weight and size, and abnormal motor neuron morphology [53–55]. Additionally, Lorson et al. [56] reported the first cloning and identification of the porcine *SMN1* gene, showing significant sequence homology between porcine and human *SMN1* in the entire coding region. Schrank et al. [57] demonstrated that the *SMN1* gene may be involved in early embryonic death. It is possible that the *SMN1* gene may influence the development of the RIBW trait by affecting proximal muscle atrophy caused by the degeneration of spinal motoneurons in pigs. However, the specific molecular mechanism underlying this relationship requires further investigation.

5. Conclusions

In this study, we performed a GWAS to investigate the TLNW trait and RIBW trait in a population of 431 DLY pigs. We identified two suggestive SNPs (ASGA0085853 and ALGA0112188) associated with TLNW, and one SNP (Affx-1115046258) associated with RIBW. Furthermore, we discovered two novel QTL regions on SSC12 (5 kb) and SSC13 (1.42 Mb) that were significantly related to TLNW and RIBW, respectively. Notably, the QTL region on SSC12 represents the first association with the TLNW trait reported to date. During further analysis, we identified three major candidate genes: *TIMP2* and *EML1* for TLNW, and *SMN1* for RIBW. This research provides valuable insights for segmenting carcass molecular breeding strategies.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ani13203243/s1>, Figure S1: Distribution of SNPs across Chromosomes after quality control for TLNW (a) and RIBW (b); Figure S2: The genotype effect plot of top SNP (Affx-115046258) related to RIBW in 408 DLY pigs (** $p < 0.01$, ns $p > 0.05$); Table S1: Distributions of SNPs after QC and the average SNPs on each chromosome of Tenderloin weight traits; Table S2: Distributions of SNPs after QC and the average SNPs on each chromosome of Ribs weight traits; Table S3: The distribution of genes within a 1 Mb range around significantly associated SNPs for the TLNW trait; Table S4: This file provides the enrichment of protein-coding genes in KEGG pathway and GO terms.

Author Contributions: Conceptualization, F.Z., J.Q., J.Y., and Z.W.; methodology, D.R. and J.Y.; software, Y.Q.; validation, R.D. and L.L.; formal analysis, C.X. and L.L.; investigation, Y.Y. and C.X.; resources, E.Z. and C.X.; data curation, G.C.; writing—original draft preparation, F.Z., J.Q., D.R., Y.Q., R.D., J.Y., Z.Z. and L.L.; writing—review and editing, J.Y., J.Q., E.Z., Y.Y., Z.Z., D.R., L.L., G.C., Y.Q. and R.D.; visualization, Z.Z., G.C., E.Z. and C.X.; supervision, L.L. and Z.Z.; project administration, Z.W. and G.C.; funding acquisition, Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Key Technologies R&D Program of Guangdong Province project (2022B0202090002) and the Project of Swine Innovation Team in the Guangdong Modern Agricultural Research System (2022KJ126).

Institutional Review Board Statement: All animals used in this study were treated in accordance with the guidelines for the use of laboratory animals set forth by the Ministry of Agriculture of China, with the approval of South China Agricultural University (Guangzhou, China), No. 2018F089 (10 August 2018).

Informed Consent Statement: Not applicable.

Data Availability Statement: The SNP genotyping data containing variant information for the DLY pigs are not publicly available because the genotyped animals belong to commercial breeding companies, but they can be obtained from the corresponding author under reasonable requirements.

Conflicts of Interest: The authors declare no conflict of interest. Zhenfang Wu and Gengyuan Cai were employed by the company Guangdong Zhongxin Breeding Technology Co., Ltd. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Font, I.F.M.; Guerrero, L. Consumer preference, behavior and perception about meat and meat products: An overview. *Meat Sci.* **2014**, *98*, 361–371. [\[CrossRef\]](#) [\[PubMed\]](#)
- da Silva, J.P.; de Alencar Nääs, I.; Abe, J.M.; da Silva Cordeiro, A.F. Classification of piglet (*Sus Scrofa*) stress conditions using vocalization pattern and applying paraconsistent logic Et. *Comput. Electron. Agric.* **2019**, *166*, 105020. [\[CrossRef\]](#)
- Lim, M.S.M.; Grohn, Y.T. Comparison of China's and the European Union's Approaches to Antimicrobial Stewardship in the Pork Industry. *Foodborne Pathog. Dis.* **2021**, *18*, 567–573. [\[CrossRef\]](#) [\[PubMed\]](#)
- Qiu, Y.; Zhuang, Z.; Meng, F.; Ruan, D.; Xu, C.; Ma, F.; Peng, L.; Ding, R.; Cai, G.; Yang, M.; et al. Identification of candidate genes associated with carcass component weights in commercial crossbred pigs through a combined GWAS approach. *J. Anim. Sci.* **2023**, *101*, skad121. [\[CrossRef\]](#)
- Zhang, L.C.; Yue, J.W.; Pu, L.; Wang, L.G.; Liu, X.; Liang, J.; Yan, H.; Zhao, K.B.; Li, N.; Shi, H.B.; et al. Genome-wide study refines the quantitative trait locus for number of ribs in a Large White × Minzhu intercross pig population and reveals a new candidate gene. *Mol. Genet. Genom.* **2016**, *291*, 1885–1890. [\[CrossRef\]](#)
- Jiang, N.; Liu, C.; Lan, T.; Zhang, Q.; Cao, Y.; Pu, G.; Niu, P.; Zhang, Z.; Li, Q.; Zhou, J.; et al. Polymorphism of VRTN Gene g.20311_20312ins291 Was Associated with the Number of Ribs, Carcass Diagonal Length and Cannon Bone Circumference in Suhuai Pigs. *Animals* **2020**, *10*, 484. [\[CrossRef\]](#)
- Kim, G.W.; Kim, H.Y. Physicochemical Quality Properties of Loin and Tenderloin Ham from Sows. *Food Sci. Anim. Resour.* **2020**, *40*, 474–483. [\[CrossRef\]](#)
- Ortiz, A.; Díaz-Caro, C.; Tejerina, D.; Escribano, M.; Crespo, E.; Gaspar, P. Consumption of fresh Iberian pork: Two-stage cluster for the identification of segments of consumers according to their habits and lifestyles. *Meat Sci.* **2021**, *173*, 108373. [\[CrossRef\]](#)
- Mikawa, S.; Hayashi, T.; Nii, M.; Shimanuki, S.; Morozumi, T.; Awata, T. Two quantitative trait loci on *Sus scrofa* chromosomes 1 and 7 affecting the number of vertebrae. *J. Anim. Sci.* **2005**, *83*, 2247–2254. [\[CrossRef\]](#)
- Uemoto, Y.; Nagamine, Y.; Kobayashi, E.; Sato, S.; Tayama, T.; Suda, Y.; Shibata, T.; Suzuki, K. Quantitative trait loci analysis on *Sus scrofa* chromosome 7 for meat production, meat quality, and carcass traits within a Duroc purebred population. *J. Anim. Sci.* **2008**, *86*, 2833–2839. [\[CrossRef\]](#)
- Ren, D.R.; Ren, J.; Ruan, G.F.; Guo, Y.M.; Wu, L.H.; Yang, G.C.; Zhou, L.H.; Li, L.; Zhang, Z.Y.; Huang, L.S. Mapping and fine mapping of quantitative trait loci for the number of vertebrae in a White Duroc × Chinese Erhualian intercross resource population. *Anim. Genet.* **2012**, *43*, 545–551. [\[CrossRef\]](#) [\[PubMed\]](#)
- Mikawa, S.; Sato, S.; Nii, M.; Morozumi, T.; Yoshioka, G.; Imaeda, N.; Yamaguchi, T.; Hayashi, T.; Awata, T. Identification of a second gene associated with variation in vertebral number in domestic pigs. *BMC Genet.* **2011**, *12*, 5. [\[CrossRef\]](#) [\[PubMed\]](#)
- Xie, L.; Qin, J.; Yao, T.; Tang, X.; Cui, D.; Chen, L.; Rao, L.; Xiao, S.; Zhang, Z.; Huang, L. Genetic dissection of 26 meat cut, meat quality and carcass traits in four pig populations. *Genet. Sel. Evol.* **2023**, *55*, 43. [\[CrossRef\]](#) [\[PubMed\]](#)
- Van Laere, A.S.; Nguyen, M.; Braunschweig, M.; Nezer, C.; Collette, C.; Moreau, L.; Archibald, A.L.; Haley, C.S.; Buys, N.; Tally, M.; et al. A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* **2003**, *425*, 832–836. [\[CrossRef\]](#)
- Oczkowicz, M.; Tyra, M.; Walinowicz, K.; Rózycki, M.; Rejdach, B. Known mutation (A3072G) in intron 3 of the IGF2 gene is associated with growth and carcass composition in Polish pig breeds. *J. Appl. Genet.* **2009**, *50*, 257–259. [\[CrossRef\]](#)
- Burgos, C.; Galve, A.; Moreno, C.; Altarriba, J.; Reina, R.; García, C.; López-Buesa, P. The effects of two alleles of IGF2 on fat content in pig carcasses and pork. *Meat Sci.* **2012**, *90*, 309–313. [\[CrossRef\]](#)
- Daetwyler, H.D.; Capitan, A.; Pausch, H.; Stothard, P.; van Binsbergen, R.; Brøndum, R.F.; Liao, X.; Djari, A.; Rodriguez, S.C.; Grohs, C.; et al. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat. Genet.* **2014**, *46*, 858–865. [\[CrossRef\]](#)
- Xu, S.S.; Gao, L.; Xie, X.L.; Ren, Y.L.; Shen, Z.Q.; Wang, F.; Shen, M.; Eyþórsdóttir, E.; Hallsson, J.H.; Kiseleva, T.; et al. Genome-Wide Association Analyses Highlight the Potential for Different Genetic Mechanisms for Litter Size Among Sheep Breeds. *Front. Genet.* **2018**, *9*, 118. [\[CrossRef\]](#)
- Zhang, H.; Shen, L.Y.; Xu, Z.C.; Kramer, L.M.; Yu, J.Q.; Zhang, X.Y.; Na, W.; Yang, L.L.; Cao, Z.P.; Luan, P.; et al. Haplotype-based genome-wide association studies for carcass and growth traits in chicken. *Poult. Sci.* **2020**, *99*, 2349–2361. [\[CrossRef\]](#)
- Fachal, L.; Dunning, A.M. From candidate gene studies to GWAS and post-GWAS analyses in breast cancer. *Curr. Opin. Genet. Dev.* **2015**, *30*, 32–41. [\[CrossRef\]](#)
- Benafif, S.; Kote-Jarai, Z.; Eeles, R.A. A Review of Prostate Cancer Genome-Wide Association Studies (GWAS). *Cancer Epidemiol. Biomark. Prev.* **2018**, *27*, 845–857. [\[CrossRef\]](#) [\[PubMed\]](#)
- Edwards, D.B.; Ernst, C.W.; Raney, N.E.; Doumit, M.E.; Hoge, M.D.; Bates, R.O. Quantitative trait locus mapping in an F2 Duroc × Pietrain resource population: II. Carcass and meat quality traits. *J. Anim. Sci.* **2008**, *86*, 254–266. [\[CrossRef\]](#) [\[PubMed\]](#)

23. Ovilo, C.; Fernández, A.; Noguera, J.L.; Barragán, C.; Letón, R.; Rodríguez, C.; Mercadé, A.; Alves, E.; Folch, J.M.; Varona, L.; et al. Fine mapping of porcine chromosome 6 QTL and LEPR effects on body composition in multiple generations of an Iberian by Landrace intercross. *Genet. Res.* **2005**, *85*, 57–67. [[CrossRef](#)] [[PubMed](#)]
24. Grover, A.; Sharma, P.C. Development and use of molecular markers: Past and present. *Crit. Rev. Biotechnol.* **2016**, *36*, 290–302. [[CrossRef](#)] [[PubMed](#)]
25. Ding, R.; Yang, M.; Wang, X.; Quan, J.; Zhuang, Z.; Zhou, S.; Li, S.; Xu, Z.; Zheng, E.; Cai, G.; et al. Genetic Architecture of Feeding Behavior and Feed Efficiency in a Duroc Pig Population. *Front. Genet.* **2018**, *9*, 220. [[CrossRef](#)]
26. Zhou, S.; Ding, R.; Meng, F.; Wang, X.; Zhuang, Z.; Quan, J.; Geng, Q.; Wu, J.; Zheng, E.; Wu, Z.; et al. A meta-analysis of genome-wide association studies for average daily gain and lean meat percentage in two Duroc pig populations. *BMC Genom.* **2021**, *22*, 12. [[CrossRef](#)]
27. Zhuang, Z.; Ding, R.; Peng, L.; Wu, J.; Ye, Y.; Zhou, S.; Wang, X.; Quan, J.; Zheng, E.; Cai, G.; et al. Genome-wide association analyses identify known and novel loci for teat number in Duroc pigs using single-locus and multi-locus models. *BMC Genom.* **2020**, *21*, 344. [[CrossRef](#)]
28. Ruan, D.; Zhuang, Z.; Ding, R.; Qiu, Y.; Zhou, S.; Wu, J.; Xu, C.; Hong, L.; Huang, S.; Zheng, E.; et al. Weighted Single-Step GWAS Identified Candidate Genes Associated with Growth Traits in a Duroc Pig Population. *Genes*. **2021**, *12*, 117. [[CrossRef](#)]
29. Chang, C.C.; Chow, C.C.; Tellier, L.C.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **2015**, *4*, 7. [[CrossRef](#)]
30. Yang, J.; Lee, S.H.; Goddard, M.E.; Visscher, P.M. GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **2011**, *88*, 76–82. [[CrossRef](#)]
31. Zhou, X.; Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **2012**, *44*, 821–824. [[CrossRef](#)] [[PubMed](#)]
32. Price, A.L.; Patterson, N.J.; Plenge, R.M.; Weinblatt, M.E.; Shadick, N.A.; Reich, D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **2006**, *38*, 904–909. [[CrossRef](#)] [[PubMed](#)]
33. Barrett, J.C.; Fry, B.; Maller, J.; Daly, M.J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **2005**, *21*, 263–265. [[CrossRef](#)]
34. Smedley, D.; Haider, S.; Ballester, B.; Holland, R.; London, D.; Thorisson, G.; Kasprzyk, A. BioMart—Biological queries made easy. *BMC Genom.* **2009**, *10*, 22. [[CrossRef](#)] [[PubMed](#)]
35. Bu, D.; Luo, H.; Huo, P.; Wang, Z.; Zhang, S.; He, Z.; Wu, Y.; Zhao, L.; Liu, J.; Guo, J.; et al. KOBAS-i: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res.* **2021**, *49*, W317–W325. [[CrossRef](#)]
36. Supek, F.; Bošnjak, M.; Škunca, N.; Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **2011**, *6*, e21800. [[CrossRef](#)]
37. Zhang, Z.; Kryvokhyzha, D.; Orsucci, M.; Glémin, S.; Milesi, P.; Lascoux, M. How broad is the selfing syndrome? Insights from convergent evolution of gene expression across species and tissues in the *Capsella* genus. *New Phytol.* **2022**, *236*, 2344–2357. [[CrossRef](#)] [[PubMed](#)]
38. Quan, J.; Ding, R.; Wang, X.; Yang, M.; Yang, Y.; Zheng, E.; Gu, T.; Cai, G.; Wu, Z.; Liu, D.; et al. Genome-wide association study reveals genetic loci and candidate genes for average daily gain in Duroc pigs. *Asian-Australas. J. Anim. Sci.* **2018**, *31*, 480–488. [[CrossRef](#)]
39. Verbeke, W.; Pérez-Cueto, F.J.; Grunert, K.G. To eat or not to eat pork, how frequently and how varied? Insights from the quantitative Q-PorkChains consumer survey in four European countries. *Meat Sci.* **2011**, *88*, 619–626. [[CrossRef](#)]
40. Grunert, K.; Wognum, N.; Trienekens, J.; Wever, M.; Olsen, N.V.; Scholderer, J. Consumer demand and quality assurance: Segmentation basis and implications for chain governance in the pork sector. *J. Chain. Netw. Sci.* **2011**, *11*, 89–97. [[CrossRef](#)]
41. Choe, J.-H.; Yang, H.-S.; Lee, S.-H.; Go, G.-W. Characteristics of pork belly consumption in South Korea and their health implication. *J. Anim. Sci. Technol.* **2015**, *57*, 22. [[CrossRef](#)] [[PubMed](#)]
42. Juárez, M.; Clemente, I.; Polvillo, O.; Molina, A. Meat quality of tenderloin from Iberian pigs as affected by breed strain and crossbreeding. *Meat Sci.* **2009**, *81*, 573–579. [[CrossRef](#)] [[PubMed](#)]
43. Fan, D.; Takawale, A.; Basu, R.; Patel, V.; Lee, J.; Kandam, V.; Wang, X.; Oudit, G.Y.; Kassiri, Z. Differential role of TIMP2 and TIMP3 in cardiac hypertrophy, fibrosis, and diastolic dysfunction. *Cardiovasc. Res.* **2014**, *103*, 268–280. [[CrossRef](#)] [[PubMed](#)]
44. Fontes, M.; Costa, T.C.; Lopes, M.M.; Souza, R.O.; Carneiro, L.S.; Paulino, P.V.R.; Chizzotti, M.L.; Silva, F.F.; Serão, N.V.L.; Duarte, M.S. Intramuscular collagen characteristics and expression of related genes in skeletal muscle of cull cows receiving a high-energy diet. *Meat Sci.* **2021**, *177*, 108495. [[CrossRef](#)]
45. Chen, Y.; Gondro, C.; Quinn, K.; Herd, R.M.; Parnell, P.F.; Vanselow, B. Global gene expression profiling reveals genes expressed differentially in cattle with high and low residual feed intake. *Anim. Genet.* **2011**, *42*, 475–490. [[CrossRef](#)]
46. Taye, M.; Kim, J.; Yoon, S.H.; Lee, W.; Hanotte, O.; Dessie, T.; Kemp, S.; Mwai, O.A.; Caetano-Anolles, K.; Cho, S.; et al. Whole genome scan reveals the genetic signature of African Ankole cattle breed and potential for higher quality beef. *BMC Genet.* **2017**, *18*, 11. [[CrossRef](#)]
47. Paganini, C.; Monti, L.; Costantini, R.; Besio, R.; Lecci, S.; Biggiogera, M.; Tian, K.; Schwartz, J.M.; Huber, C.; Cormier-Daire, V.; et al. Calcium activated nucleotidase 1 (CANT1) is critical for glycosaminoglycan biosynthesis in cartilage and endochondral ossification. *Matrix Biol.* **2019**, *81*, 70–90. [[CrossRef](#)]

48. Oegema, R.; McGillivray, G.; Leventer, R.; Le Moing, A.G.; Bahi-Buisson, N.; Barnicoat, A.; Mandelstam, S.; Francis, D.; Francis, F.; Mancini, G.M.S.; et al. EML1-associated brain overgrowth syndrome with ribbon-like heterotopia. *Am. J. Med. Genet. C Semin. Med. Genet.* **2019**, *181*, 627–637. [[CrossRef](#)]
49. Collin, G.B.; Won, J.; Krebs, M.P.; Hicks, W.J.; Charette, J.R.; Naggert, J.K.; Nishina, P.M. Disruption in murine Eml1 perturbs retinal lamination during early development. *Sci. Rep.* **2020**, *10*, 5647. [[CrossRef](#)]
50. Affar, E.B.; Gay, F.; Shi, Y.; Liu, H.; Huarte, M.; Wu, S.; Collins, T.; Li, E.; Shi, Y. Essential dosage-dependent functions of the transcription factor yin yang 1 in late embryonic development and cell cycle progression. *Mol. Cell. Biol.* **2006**, *26*, 3565–3581. [[CrossRef](#)]
51. Wirth, B. An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Hum. Mutat.* **2000**, *15*, 228–237. [[CrossRef](#)]
52. Poirier, A.; Weetall, M.; Heinig, K.; Bucheli, F.; Schoenlein, K.; Alsenz, J.; Bassett, S.; Ullah, M.; Senn, C.; Ratni, H.; et al. Risdiplam distributes and increases SMN protein in both the central nervous system and peripheral organs. *Pharmacol. Res. Perspect.* **2018**, *6*, e00447. [[CrossRef](#)] [[PubMed](#)]
53. Gogliotti, R.G.; Quinlan, K.A.; Barlow, C.B.; Heier, C.R.; Heckman, C.J.; Didonato, C.J. Motor neuron rescue in spinal muscular atrophy mice demonstrates that sensory-motor defects are a consequence, not a cause, of motor neuron dysfunction. *J. Neurosci.* **2012**, *32*, 3818–3829. [[CrossRef](#)]
54. Bowerman, M.; Anderson, C.L.; Beauvais, A.; Boyl, P.P.; Witke, W.; Kothary, R. SMN, profilin IIa and plastin 3: A link between the deregulation of actin dynamics and SMA pathogenesis. *Mol. Cell Neurosci.* **2009**, *42*, 66–74. [[CrossRef](#)] [[PubMed](#)]
55. Park, G.H.; Maeno-Hikichi, Y.; Awano, T.; Landmesser, L.T.; Monani, U.R. Reduced survival of motor neuron (SMN) protein in motor neuronal progenitors functions cell autonomously to cause spinal muscular atrophy in model mice expressing the human centromeric (SMN2) gene. *J. Neurosci.* **2010**, *30*, 12005–12019. [[CrossRef](#)]
56. Lorson, M.A.; Spate, L.D.; Prather, R.S.; Lorson, C.L. Identification and characterization of the porcine (*Sus scrofa*) survival motor neuron (SMN1) gene: An animal model for therapeutic studies. *Dev. Dyn.* **2008**, *237*, 2268–2278. [[CrossRef](#)]
57. Schrank, B.; Götz, R.; Gunnensen, J.M.; Ure, J.M.; Toyka, K.V.; Smith, A.G.; Sendtner, M. Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 9920–9925. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Identification of candidate genes associated with carcass component weights in commercial crossbred pigs through a combined GWAS approach

Yibin Qiu,^{†,1} Zhanwei Zhuang,^{†,1} Fanming Meng,^{†,1} Donglin Ruan,[†] Cineng Xu,[†] Fucai Ma,[†] Longlong Peng,[†] Rongrong Ding,^{†,||,§} Gengyuan Cai,^{†,||,§} Ming Yang,^{**} Zhenfang Wu,^{†,§,¶} Jie Yang,^{†,§,2} and Enqin Zheng^{†,§,2}

[†]College of Animal Science and National Engineering Research Center for Breeding Swine Industry, South China Agricultural University, Guangzhou 510642, PR China

[‡]State Key Laboratory of Livestock and Poultry Breeding/Guangdong Key Laboratory of Animal Breeding and Nutrition, Institute of Animal Science, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, PR China

^{||}Guangdong Zhongxin Breeding Technology Co., Ltd, Guangzhou 511400, PR. China

[§]Guangdong Provincial Key Laboratory of Agro-animal Genomics and Molecular Breeding, South China Agricultural University, Guangzhou 510642, PR China

[¶]Yunfu Subcenter of Guangdong Laboratory for Lingnan Modern Agriculture, Yunfu 527400, PR China

^{**}College of Animal Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, PR China

¹These authors contributed equally to this work.

²Corresponding author: jieyang2012@hotmail.com

Abstract

In the pork industry chain, carcass cutting is crucial for enhancing the commercial value of pork carcasses. However, the genetic mechanisms underlying carcass component weights remain poorly understood. Here, we used a combined genome-wide association study (GWAS) approach that integrated single- and multi-locus models to map genetic markers and genes associated with the weights of seven carcass components in Duroc × Landrace × Yorkshire (DLY) pigs. As multi-locus GWAS captures more single nucleotide polymorphisms (SNPs) with large effects than single-locus GWAS, the combined GWAS approach detected more SNPs than using the single-locus model alone. We identified 177 nonredundant SNPs associated with these traits in 526 DLY pigs, including boneless butt shoulder (BBS), boneless picnic shoulder (BPS), boneless leg (BL), belly (BELLY), front fat (FF), rear fat (RF), and skin-on whole loin (SLOIN). Using single-locus GWAS, we identified a quantitative trait locus (QTL) for SLOIN on *Sus scrofa* chromosome 15 (SSC15). Notably, a single SNP (ASGA0069883) in the proximity of this QTL was consistently detected by all GWAS models (one single-locus and four multi-locus models) and explained more than 4% of the phenotypic variance. Our findings suggest that the involved gene, *MYO3B*, is proposed to be a strong candidate for SLOIN. Further analysis also identified several candidate genes related to BBS (*PPP3CA* and *CPEB4*), BPS (*ECH1*), FF (*CACNB2* and *ZNF217*), BELLY (*FGFRL1*), BL (*CHST11*), and RF (*LRRK2*). The identified SNPs can be used as molecular markers for the genetic improvement of pork carcasses in the molecular-guided breeding of modern commercial pigs.

Lay Summary

Carcass cutting is the most effective method for enhancing the commercial value of pork carcasses in the industry chain. However, the genetic mechanisms underlying carcass component weights remain elusive. In this study, we used a combination of single- and multi-locus models to increase the power of genome-wide association analysis. We identified 177 important genetic variants that are potentially promising candidate markers for marker-assisted selection in breeding. Further investigation revealed one quantitative trait locus region and several candidate genes (*PPP3CA*, *CPEB4*, *ECH1*, *CACNB2*, *ZNF217*, *FGFRL1*, *CHST11*, *LRRK2*) associated with the weights of seven carcass components in Duroc × Landrace × Yorkshire pigs.

Key words: carcass component weights, genome-wide association study, multi-locus, pigs, single-locus

Abbreviations: BBS, boneless butt shoulder; BELLY, belly; BL, boneless leg; BPS, boneless picnic shoulder; FF, front fat; GWAS, genome-wide association study; LD, linkage disequilibrium; MLM, mixed linear model; QTL, quantitative trait locus; RF, rear fat; SLOIN, skin-on whole loin; SNP, single nucleotide polymorphism; SSC, *sus scrofa* chromosome

Introduction

Pork quality is primarily determined by the palatability and nutrition of the meat, which can vary depending on the cut of pork (Alfaia et al., 2019; Chen et al., 2021). In this case,

some parts of the meat will be favored, and the price will be higher than others. Therefore, in the pork industry chain, carcass cutting is the most effective method to improve the commercial value of pork carcasses. Previous research has

Received December 28, 2022 Accepted April 21, 2023.

© The Author(s) 2023. Published by Oxford University Press on behalf of the American Society of Animal Science. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

identified several quantitative trait loci (QTLs) and candidate genes that underlie carcass traits in pigs, shedding light on the genetic basis of these economically important traits. For example, a QTL on *Sus scrofa* chromosome 1 (SSC1) that affects the carcass quality in pigs was identified in 1998 using a Meishan × White reciprocal backcross population (Rohrer and Keele, 1998). The *VRTN* gene has also been shown to indirectly contribute to carcass length by affecting swine vertebral numbers (Mikawa et al., 2011; Yang et al., 2016). Lee et al. showed that a genetic variation of *CRTC3* can be used to predict the amount of belly fat, emphasizing the importance of genetic factors in determining carcass traits (Lee et al., 2018b). Furthermore, Palombo et al. have identified several QTL signals for untrimmed and trimmed thigh weight in Italian crossbred pigs (Palombo et al., 2021). However, the genetic mechanism responsible for carcass traits remains elusive due to the diversity of available products, suggesting that there is still considerable potential for further genetic improvement in carcass traits.

Genome-wide association study (GWAS) is a practical approach to the genetic dissection of quantitative traits that has been widely used in livestock (Ding et al., 2019; Zhuang et al., 2020; Zhou et al., 2021). More recently, several studies in pigs have shown that GWAS enables researchers to efficiently uncover quantitative trait genes or nucleotides (QTNs) with complex traits (Ma et al., 2014; Cho et al., 2019; Nosková et al., 2020). However, most complex traits are controlled by multiple loci, and single-locus GWAS may not recover the true model due to the intrinsic limitation of the model (Wang et al., 2016), resulting in “missing heritability” (Maher, 2008; Ding et al., 2019). Additionally, the typical Bonferroni correction tends to be conservative, so many vital loci may fail to pass the stringent criterion of a significance test. Several multi-locus models have been proposed to overcome these bottlenecks, which have higher power for detecting QTN by addressing the locus heterogeneity (Galvan et al., 2010) compared to single-locus GWAS analysis (Segura et al., 2012; Wang et al., 2016). For example, the multi-locus random-single nucleotide polymorphism (SNP)-effect mixed linear model (mrMLM; Wang et al., 2016), fast multi-locus random-SNP-effect mixed linear model (FASTmrMLM; Tamba and Zhang, 2018; Zhang et al., 2020), fast multi-locus random-SNP-effect efficient mixed model association (FASTmrEMMA; Wen et al., 2017), and pKwMEB (Ren et al., 2018). These methods effectively control the false positive rate and reserve more significant loci for complex traits (Wang et al., 2016; Wen et al., 2017; Zhang et al., 2020).

To further unearth loci associated with carcass component weights, we performed a combined GWAS approach (using both single-locus and multi-locus analysis) on 526 Duroc × Landrace × Yorkshire (DLY) pigs, one of the most common crossbred pigs in the world. The objective of this study was to dissect the genetic structure of carcass traits and facilitate the rapid development of molecular breeding in pigs.

Materials and Methods

Animals and phenotype

The animals and experimental methods used in this study follow the guidelines of the Ministry of Agriculture of China and the Use Committee of South China Agricultural University (SCAU). The ethics committee of SCAU (Guangzhou, China) approved all animal experiments.

From 2018 to 2019, we collected 526 three-way crossbred DLY pigs intercrossed by Duroc boars and (Landrace × Yorkshire) sows from Wen's Foodstuffs Group Co., Ltd (Guangdong, China). All pigs were raised under the same management condition on the same farm, and provided with ad libitum access to drinking water and feed. The slaughtering procedure followed the protocol from our previous study (Zhang et al., 2021), with the pigs being slaughtered in five batches at a commercial abattoir in Chifeng, Inner Mongolia, at approximately 210 d of age. The carcass was sectioned into anterior, middle, and posterior portions by sawing between the 5th and 6th ribs and at the lumbosacral junction. The left half of the carcass components were finely cut and weighed within 30 min postmortem (Zhang et al., 2021). We recorded the weights of the following carcass components for all pigs: boneless butt shoulder (BBS), boneless picnic shoulder (BPS), boneless leg (BL), belly (BELLY), front fat (FF), rear fat (RF), and skin-on whole loin (SLOIN). In this study, all meat was deboned, skinned, and trimmed of external fat, except for SLOIN, which remained unskinned. Further details on the cutting position are shown in Figure 1A and Supplementary Table S1.

Genotyping and quality control

Genotyping was performed as described by our previous studies (Ding et al., 2019; Qiu et al., 2021). Briefly, Genomic DNA was extracted from ear tissue samples, and the quality of DNA in all samples (526 DNA samples) was assessed by the ratios of light absorption (A260/280 and A260/230) and gel electrophoresis, using a DNA concentration of 50 ng/μL. Samples were genotyped with the GeneSeek Porcine 50K SNP Chip (Neogen, Lincoln, NE, United States). Quality control was performed using the PLINK software v1.90 (Chang et al., 2015). Individuals with an overall call rate < 95% were removed. SNP with call rates < 90%, minor allele frequency < 1%, and Hardy-Weinberg *P*-value < 10⁻⁶ were eliminated. Moreover, all SNPs located on the sex chromosomes and unmapped regions were also excluded. After filtering, a final set of 35,235 SNPs from 526 DLY pigs remained for further statistical analyses.

Single-locus GWAS

Single-locus GWAS of seven traits were performed on GEMMA software v0.98.1 (Zhou and Stephens, 2012), using a univariate mixed linear model (MLM) to test the association between each SNP and phenotype. The MLM was as follows:

$$y = W\alpha + X\beta + u + \varepsilon$$

where *y* represents a vector of the phenotypes in the DLY pigs; *W* is the incidence matrix of covariates, including fixed effects of sex, slaughter batch, live weight, and the top five eigenvectors obtained prior to this analysis using the GCTA software v1.92.4beta (Yang et al., 2011); *α* represents the vector of corresponding coefficients including the intercept; *X* is the vector of SNP genotypes; *β* specifies the corresponding effect size of the marker; *u* is the vector of random effects, with *u* ~ MVN_{*n*} (0, λτ⁻¹*K*); *ε* is the vector of random residuals, with *ε* ~ MVN_{*n*} (0, τ⁻¹*I*_{*n*}); λ is the ratio between the two variance components; τ⁻¹ is the variance of the residual errors; *K* is a genomic relatedness matrix (GRM) estimated by the GEMMA software v0.98.1; *I* is an *n* × *n*

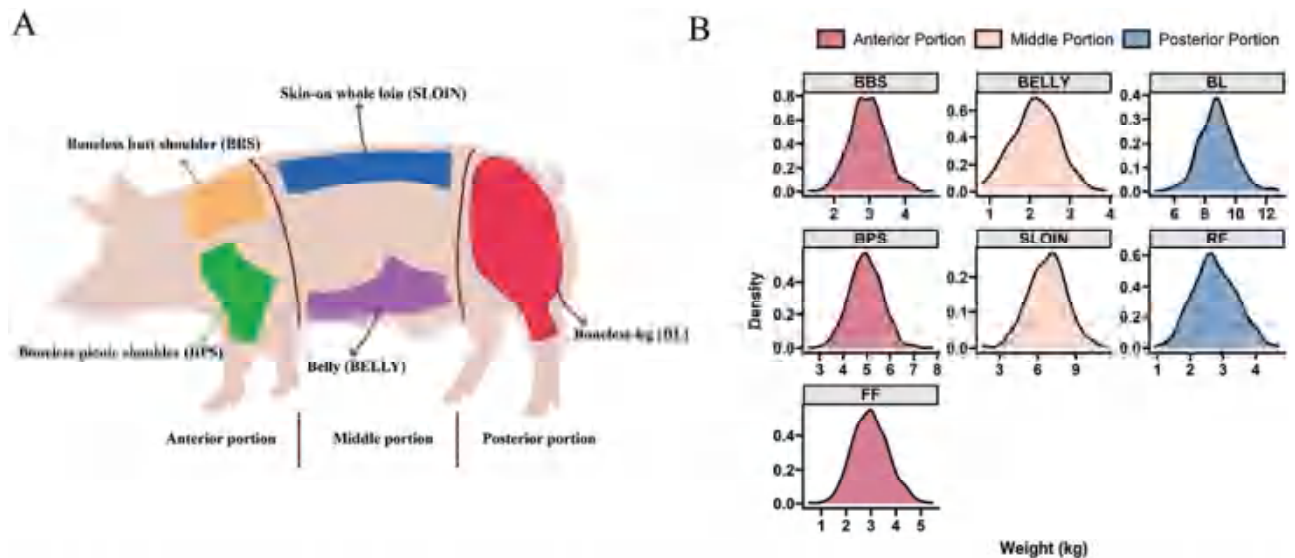


Figure 1. Schematic diagram of carcass cutting after pig slaughter (A) and distribution of seven carcass traits in 526 DLY pigs (B). The anterior portion includes BBS (boneless butt shoulder), BPS (boneless picnic shoulder), and FF (front fat). The middle portion includes BELL (belly) and SLOIN (skin-on whole loin). The posterior portion includes BL (boneless leg) and RF (rear fat). FF and RF are not labeled in (A).

identity matrix; MVN_n denotes the n -dimensional multivariate normal distribution. The suggestive and genome-wide significance thresholds were determined by the Bonferroni method (Yang et al., 2005). We also computed the genomic control λ (λ_{GC}) to evaluate whether confounding due to population stratification exists. The λ_{GC} is defined as the median χ^2 (1 degree of freedom) association statistic across SNPs divided by its theoretical median under the null distribution (Devlin and Roeder, 1999; Pritchard and Rosenberg, 1999; Reich and Goldstein, 2001).

Haplotype block analysis and conditional analysis

The PLINK v1.90 (Chang et al., 2015) and Haploview v4.2 (Barrett et al., 2005) were used for chromosomal regions with multiple significantly clustered around the top SNP to evaluate the linkage disequilibrium (LD) pattern of the region. LD blocks were defined according to Gabriel et al. (Gabriel et al., 2002). GWAS often identifies a significant set of SNPs associated with target traits in putative regions, possibly due to high LD between them. To distinguish the independence of all significant signals in putative regions, we performed a conditional analysis in MLM by fitting the genotypes of the top SNP as covariates (Yang et al., 2012).

Multi-locus GWAS

Four multi-locus GWAS methods were performed on seven traits in the R package “mrMLM” v4.0.2 (Zhang et al., 2020). The models used included mrMLM (Wang et al., 2016), FASTmrMLM (Tamba and Zhang, 2018; Zhang et al., 2020), FASTmrEMMA (Wen et al., 2017), and pKWmEB (Ren et al., 2018). These models consist of two steps. All potentially associated markers are selected in the first stage using various algorithms. In the following stage, the selected SNPs are included in the multi-locus model. Then, their effects are estimated using empirical Bayes, and all nonzero effects are further assessed with the likelihood ratio test for true QTNs (Zhang et al., 2020). The population genetic structure was the same as that used in single-locus GWAS, and GRM was calculated by the R package “mrMLM” v4.0.2 (Zhang

et al., 2020). Other parameters were set to default values, as described by Ding et al. (Ding et al., 2019) and Zhuang et al. (Zhuang et al., 2020). The markers surpassing the log of odds (LOD) score threshold (default value: 3) were considered promising trait-associated SNPs (Li et al., 2018; Ding et al., 2019; Zhang et al., 2020; Zhuang et al., 2020).

Genetic parameters and heritability estimation

In this study, we calculated the phenotypic correlation of the sample via Pearson's correlation coefficient. We also used the GCTA software v1.92.4beta (Yang et al., 2011) to estimate the genetic correlation between two traits by performing a bivariate genome-based restricted maximum likelihood analysis. The formula was as follows (Yang et al., 2011; Deary et al., 2012; Lee et al., 2012):

$$r_g = \frac{\sigma_{g1g2}}{\sigma_{g1}\sigma_{g2}}$$

where r_g is the genetic correlation coefficient between two traits; the subscripts ‘1’ and ‘2’ represent the two traits; σ_{g1g2} refers to the genetic covariance between traits; σ_g represents the square root of the genetic variance for each trait (captured by the SNPs).

GCTA software v1.92.4beta (Yang et al., 2011) was also utilized to estimate the heritability contributed by the genome-wide SNPs, as well as the significant SNPs identified through single- and multi-locus GWAS, in the bivariate restricted maximum likelihood model. The formula was as follows (Yang et al., 2011):

$$y = X\beta + \sum_{i=1}^r g_i + \varepsilon \text{ with } \text{var}(y) = \sum_{i=1}^r A_i \sigma_i^2 + I\sigma_\varepsilon^2$$

where y is the vector of phenotypes; β is the vector of fixed effects; X refers to an incidence matrix for β ; r refers to either the number of genome-wide SNPs or the number of selected SNPs; g_i is the vector of the genetic effects attributed to the

i^{th} SNP; σ_i^2 is the variance of i^{th} genetic factor with its corresponding GRM, A_i ; I is the identity matrix, and σ_e^2 refers to the residual variance.

Candidate gene identification and functional enrichment analysis

All SNPs refer to the latest version of the *Sus scrofa* 11.1 genome (http://ensembl.org/Sus_scrofa/Info/Index). Functional gene annotations (v105) were downloaded in GFF3 format from the Ensembl website (http://ftp.ensembl.org/pub/release-105/gff3/sus_scrofa/). As porcine gene annotations are incomplete, we converted pig Ensembl gene IDs to orthologous human Ensembl gene IDs using BioMart (<http://asia.ensembl.org/biomart/martview>). To gain insight into potential candidate genes, we performed Gene Ontology (GO) term annotation, National Human Genome Research Institute (NHGRI) GWAS catalog annotation, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis using KOBAS v3.0 (Bu et al., 2021). Enriched terms with corrected P -value < 0.05, as determined by Fisher's exact test and Benjamini-Hochberg correction (Benjamini and Hochberg, 1995), were selected for further exploration of genes involved in biological pathways and processes. GeneCards (<http://www.genecards.org/>) and Ensembl (www.ensembl.org/biomart/martview) were used to query gene functions.

Results

Phenotype and heritability statistics

In this study, we recorded the weights of seven carcass components in 526 DLY pigs and summarized the statistics for the anterior (BBS, BPS, and FF), middle (BELLY and SLOIN), and posterior (BL and RF) portions in Table 1. Before GWAS analysis, we assessed the distribution of all phenotypes using the Kolmogorov–Smirnov test (Marsaglia et al., 2003) and confirmed that all phenotypic data followed a Gaussian distribution (Figure 1B). We observed great phenotypic diversity within the population. For example, the FF weight ranged from 0.52 to 5.46 kg, BELLY ranged from 0.82 to 3.88 kg, and SLOIN ranged from 1.66 to 11.31 kg (Table 1; Figure 1B). The coefficient of variation for each trait ranged from 13.23% to 26.06%, also indicating substantial phenotypic variation

(Table 1). Supplementary Figure S1 shows that the traits were highly positively correlated, as indicated by the phenotypic and genetic correlations. We also estimated the genomic heritability for each trait, which varied widely and ranged from 0.16 to 0.55 (Table 1). Traits with greater heritability values included BPS (0.55) and SLOIN (0.47), suggesting that the genomic variants could explain substantial proportions of the phenotypic variation for these traits.

Single-locus GWAS results

After a series of quality control procedures, 35,235 SNPs with genotypes on 526 DLY pigs were retained for GWAS analysis. According to the number of SNPs used for single-locus GWAS, the genome-wide and suggestive thresholds were 1.42×10^{-6} (0.05/35,235) and 2.84×10^{-5} (1/35,235), respectively. The GWAS λ_{GC} ranged from 1.00 to 1.05, indicating that the population stratification was adequately adjusted (Supplementary Figure S2; Price et al., 2010).

The results of the single-locus GWAS are shown in Table 2, Figures 2 and 3. A total of 19 SNPs, including 16 suggestive and 3 genome-wide significant SNPs, were identified as associated with the weights of seven carcass components. In the anterior portion, we identified two suggestive SNPs associated with BBS, one suggestive and one genome-wide significant SNP associated with BPS, and one suggestive SNP associated with FF. In the middle portion, one suggestive SNP was associated with BELLY, and six SNPs (four suggestive and two genome-wide significant) were associated with SLOIN. In the posterior portion, we also identified five suggestive SNPs associated with BL and two suggestive SNPs associated with RF.

We observed two lead SNPs (H3GA0044562 and H3GA0044566) in complete LD and located within a 115 kb haplotype block (15: 77,542,253–77,657,791 bp), which suggests that mutations near this QTL region may have a substantial effect on SLOIN (Figure 4C). The top markers for the above-characterized haplotype blocks accounted for 4.20% of the phenotypic variance for SLOIN (Table 2). Among them, carriers of the G allele of H3GA0044562 (GG and AG) had greater SLOIN weight than noncarriers (Figure 4D). To further investigate, we conducted a conditional analysis by fitting H3GA0044562 as a covariate into a univariate MLM using GEMMA (Zhou and Stephens, 2012; Figure 4A and B).

Table 1. Phenotype and heritability statistics for carcass traits

Portion	Trait ¹	N ²	Unit	Mean (SD) ³	Min ⁴	Max ⁵	C.V. (%) ⁶	h^2 (SE) ⁷
Anterior	BBS	493	kg	2.95 ± 0.49	1.31	4.76	16.61	0.28 ± 0.09
	BPS	493	kg	4.96 ± 0.7	2.56	7.79	14.11	0.55 ± 0.08
	FF	493	kg	2.98 ± 0.71	0.52	5.46	23.83	0.28 ± 0.08
Middle	BELLY	416	kg	2.15 ± 0.56	0.82	3.88	26.05	0.34 ± 0.10
	SLOIN	418	kg	6.77 ± 1.48	1.66	11.31	21.86	0.47 ± 0.10
Posterior	BL	505	kg	8.77 ± 1.16	4.78	12.8	13.23	0.35 ± 0.09
	RF	486	kg	2.76 ± 0.66	0.94	4.7	23.91	0.16 ± 0.08

¹BBS, boneless butt shoulder; BPS, boneless picnic shoulder; BL, boneless leg; BELLY, belly; FF, front fat; RF, rear fat; SLOIN, skin-on whole loin.

²N, Number.

³Mean (standard deviation).

⁴Min, minimum.

⁵Max, maximum.

⁶C.V., coefficient of variation.

⁷ h^2 (SE), Heritability (standard error).

Table 2. Description of significant SNPs identified by single-locus GWAS

Portion	Trait ¹	SSC ²	SNP ID	Position (bp) ³	P-value ⁴	EPV (%) ⁵	Nearest gene	Distance (bp) ⁶
Anterior	BBS	8	ASGA0039766	119,699,034	1.74E-05	0.52%	PPP3CA	within
	BBS	10	WU_10.2_10_74855453	68,144,930	2.49E-05	0.64%	ADARB2	3,157
	BPS	6	Hal_2	47,357,966	1.43E-07	2.44%	ENSSSCG00000041601	6,238
	BPS	12	WU_10.2_12_4237185	4,490,123	9.17E-06	2.14%	SEC14L1	-1,043
	FF	7	WU_10.2_7_16120808	15,224,170	2.83E-05	0.15%	ID4	57,451
Middle	BELLY	2	ASGA0011472	117,990,345	7.22E-06	2.07%	YTHDC2	within
	SLOIN	1	DRGA0002342	252,300,956	1.78E-05	3.54%	ECPAS	within
	SLOIN	15	ASGA0069883	76,663,020	4.42E-06	4.57%	MYO3B	within
	SLOIN	15	ALGA0085978	76,994,963	3.47E-06	3.36%	GAD1	within
	SLOIN	15	MARC0046707	77,157,077	2.82E-05	2.53%	TLK1	13,009
	SLOIN	15	H3GA0044562	77,566,996	7.75E-07	4.20%	DCAF17	within
Posterior	SLOIN	15	H3GA0044566	77,643,035	7.75E-07	4.20%	CYBRD1	6,005
	BL	1	DRGA0002413	258,764,579	2.82E-05	1.91%	ENSSSCG00000050756	-7,019
	BL	3	ALGA0017416	10,594,633	6.33E-06	5.89%	ENSSSCG00000007705	15,222
	BL	3	DBWU0000145	10,657,335	8.47E-06	6.07%	TRIM50	-997
	BL	3	ASGA0013321	10,931,446	5.95E-06	5.55%	MLXIPL	within
	BL	6	WU_10.2_6_30616257	34,906,868	2.14E-05	2.21%	ZNF423	25,726
	RF	3	WU_10.2_3_4803085	4,160,760	4.83E-06	4.85%	RNF216	within
	RF	11	WU_10.2_11_46686334	42,441,803	1.92E-05	0.69%	KLHL1	-22,033

¹BBS, boneless butt shoulder; BPS, boneless picnic shoulder; BL, boneless leg; BELLY, belly; FF, front fat; RF, rear fat; SLOIN, skin-on whole loin.

²SSC, *Sus scrofa* chromosome.

³SNP position in Ensembl.

⁴P-value in boldface: genome-wide significant.

⁵EPV, explained phenotypic variance.

⁶+/-: the SNP located upstream/downstream of the nearest gene.

The *P*-values association with SLOIN fell below the suggested threshold, indicating the putative SNPs had a high LD (Ding et al., 2021; Zhou et al., 2021).

Multi-locus GWAS results

Multi-locus GWAS has the advantage of detecting more genetic variants than single-locus GWAS (Zhang et al., 2020). Therefore, we also performed multi-locus GWAS for the seven traits using various models, including mrMLM (Wang et al., 2016), FASTmrMLM (Tamba and Zhang, 2018; Zhang et al., 2020), FASTmrEMMA (Wen et al., 2017), and pKW-mEB (Ren et al., 2018).

In the DLY population, we identified a total of 23, 30, 26, 25, 13, 27, and 26 significant SNPs (LOD score > 3) for BBS, BPS, FF, BELLY, SLOIN, BL, and RF, respectively, using the aforementioned models (Supplementary Table S2, Figures S3 and S4). In further analysis, one SNP on SSC15 (ASGA0069883), identified by single-locus GWAS, was also found by four multi-locus models for SLOIN (Table 2, Supplementary Table S2, and Supplementary Figure S5). This SNP explained more than 4% of the phenotypic variance (in all models) and was close to the leading SNPs detected by single-locus GWAS (H3GA0044562 and H3GA0044566). The findings suggest that the involved gene, *MYO3B*, is proposed to be a strong candidate for SLOIN.

Capturing missing heritability

In addition to the SNPs identified by single-locus GWAS, multi-locus GWAS identified more significant SNPs on other chromosomes (Supplementary Figure S5). We separately com-

pared the heritability contributed by the genome-wide SNPs, as well as the significant SNPs identified through single- and multi-locus GWAS, for each trait (Figure 5 and Supplementary Table S3). Our observation shows that significant SNPs identified by multi-locus GWAS contribute a larger proportion of heritability than single-locus GWAS (Wilcoxon rank sum test, $P = 5.28 \times 10^{-4}$). This result suggests that the ability of multi-locus GWAS to capture missing heritability might be mainly due to the effective inclusion of more large-effect SNPs (Supplementary Tables S2 and S3; Wang et al., 2016; Wen et al., 2017; Ren et al., 2018; Tamba and Zhang, 2018; Zhang et al., 2020; Zhuang et al., 2020).

Candidate genes search and functional annotation

We conducted LD decay analysis in the DLY population and found that LD (r^2) dropped below 0.2 after a distance greater than 150 kb (Supplementary Figure S6). This suggests that the causal mutations and genes are likely to be located within a region spanning 500 kb upstream and downstream of the identified GWAS signals. Based on the annotations of the *Sus scrofa* 11.1 genome assembly, we found 1291 protein-coding genes within a 1 Mb region centered on the 177 nonredundant GWAS signals. Our pathway enrichment analysis revealed several significantly enriched KEGG pathways and GO terms related to carcass traits, including metabolic pathways and protein binding (Supplementary Table S4 and Figure 6). Annotation results from the NHGRI GWAS catalog further suggested that most of these genes appear to have physiological roles in obesity-related traits (Supplementary Table S4 and Figure 6). Additionally, we used the GeneCards database and literature

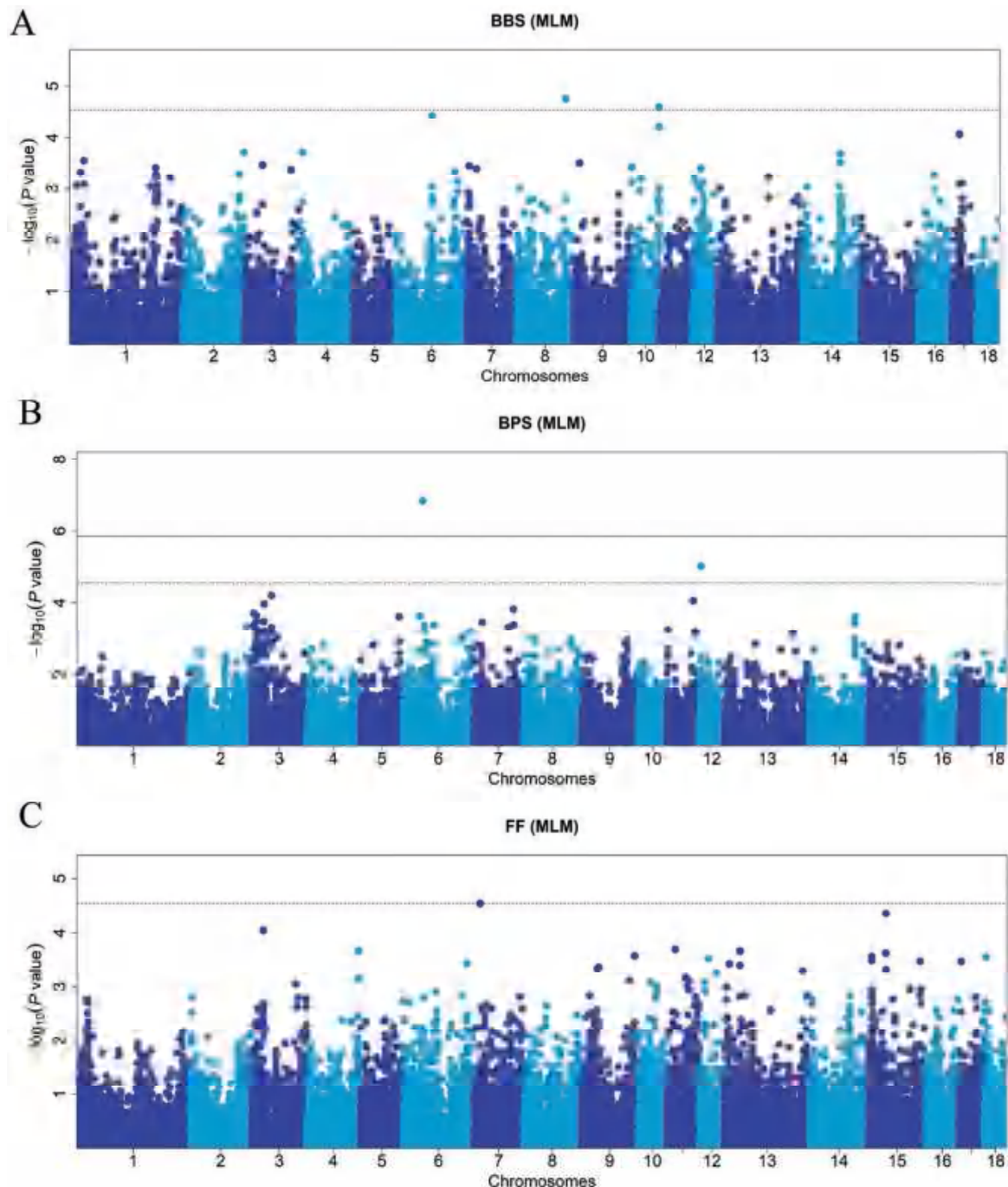


Figure 2. Manhattan plots of the single-locus GWAS for anterior carcass traits in 526 DLY pigs. Manhattan plots consisted of (A) BBS (boneless butt shoulder), (B) BPS (boneless picnic shoulder), and (C) FF (front fat). The x-axis represents the chromosomes, and the y-axis represents the $-\log_{10}(P\text{-value})$. The solid and dashed lines indicate the 5% genome-wide (1.42×10^{-6}) and suggestive (2.84×10^{-5}) Bonferroni-corrected thresholds, respectively.

to investigate the functions of these genes involved in the high-lighted terms. Finally, nine functional genes were tentatively considered as candidates related to BBS (*PPP3CA* and *CPEB4*), BPS (*ECH1*), FF (*CACNB2* and *ZNF217*), BELLY (*FGFRL1*), SLOIN (*MYO3B*), BL (*CHST11*), and RF (*LRRK2*).

Discussion

The DLY pigs are the most widely used crossbreed in commercial pig production (Li et al., 2021). Dissecting the genetic basis of carcass traits in DLY pigs will promote carcass qual-

ity and yield, allowing producers to achieve greater profitability throughout the chain.

Combining single- and multi-locus GWASs can improve GWAS discovery efficiency

Complex traits in pigs are often controlled by multiple loci (Wu et al., 2021). However, traditional single-locus GWAS can only identify a small fraction of genetic variation. By comparing the heritability captured by SNPs identified in single- and multi-locus GWAS, our study revealed that multi-locus GWAS identified more SNPs with a large effect (Figure 5,

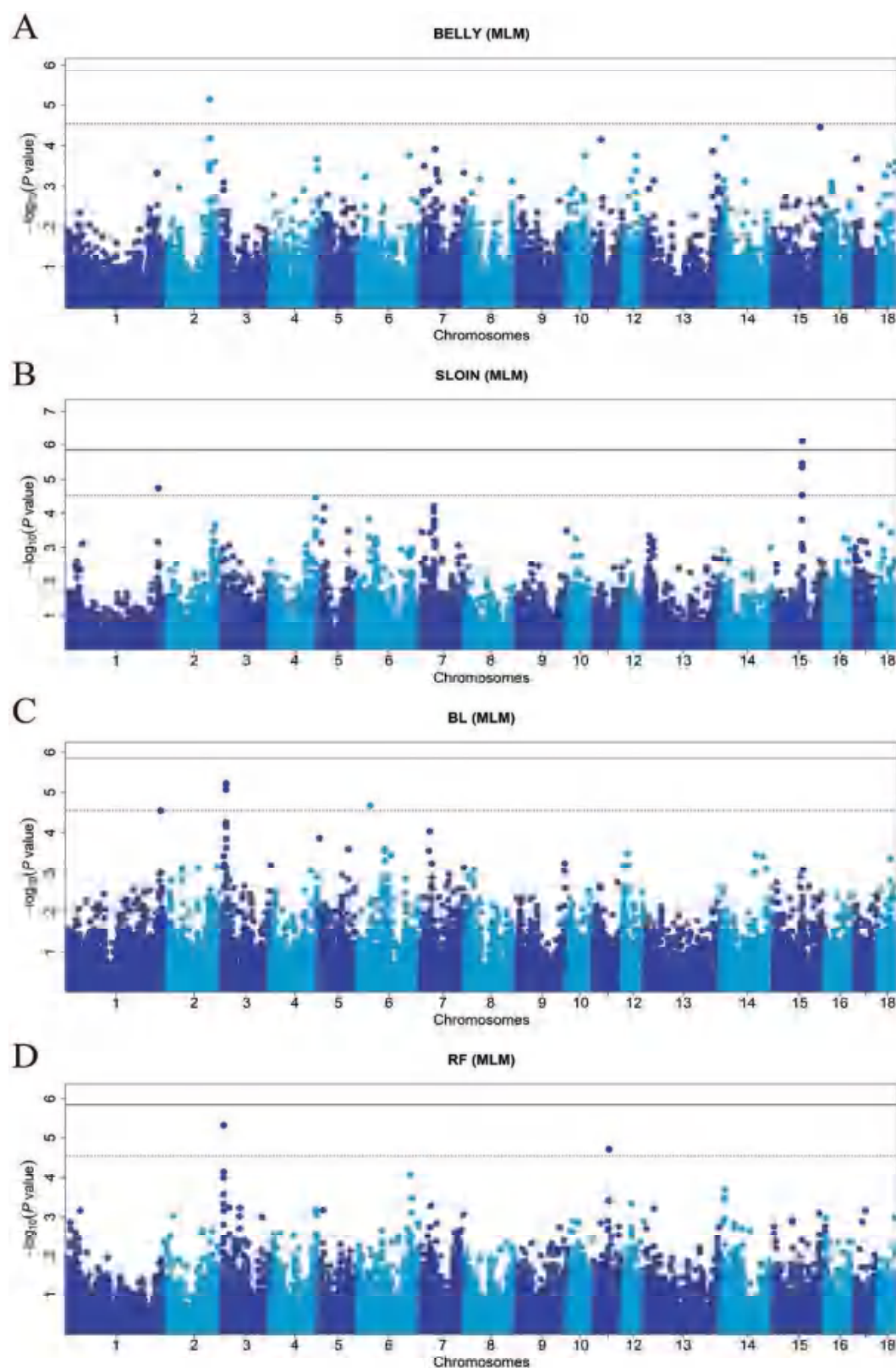


Figure 3. Manhattan plots of the single-locus GWAS for middle and posterior carcass traits in 526 DLY pigs. Manhattan plots consisted of (A) BELLY (belly), (B) SLOIN (skin-on whole loin), (C) BL (boneless leg), and (D) RF (rear fat). The x-axis represents the chromosomes, and the y-axis represents the $-\log_{10}(P\text{-value})$. The solid and dashed lines indicate the 5% genome-wide (1.42×10^{-6}) and suggestive (2.84×10^{-5}) Bonferroni-corrected thresholds, respectively.

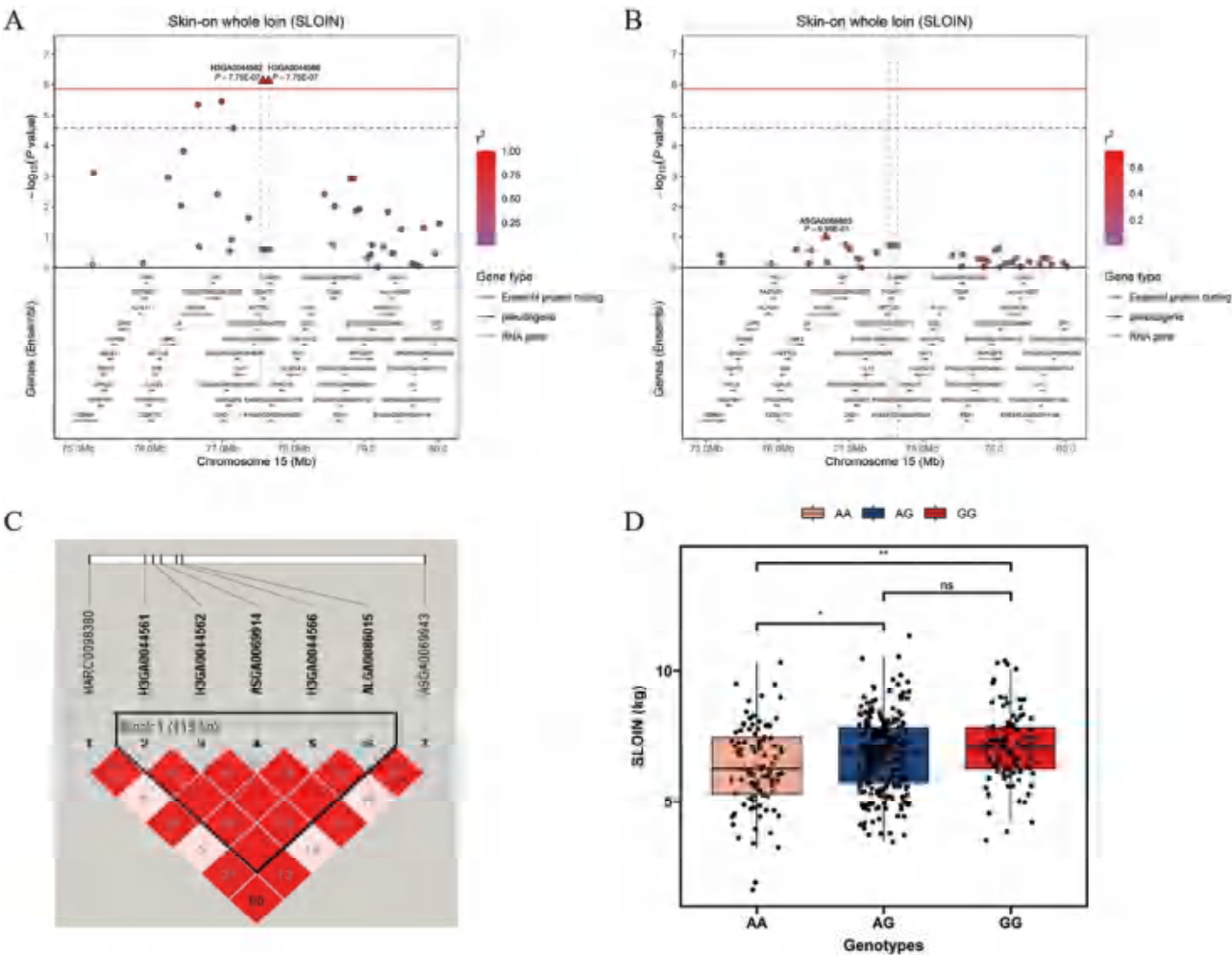


Figure 4. Regional association plots of the significant single nucleotide polymorphisms (SNPs) associated with SLOIN at SSC15. For (A) and (B), negative log₁₀(P-value) of SNPs (y-axis) are presented according to their chromosomal positions (x-axis). The red line and black dashed line indicate the 5% genome-wide significance level (1.42×10^{-6}) and the suggestive level (2.84×10^{-5}), respectively. The significant top SNPs are indicated by big triangles. SNPs are denoted by color shades depending on the target SNP with which they were in the strongest linkage disequilibrium. The plots of (A) and (B) show the association results for SLOIN before and after conditional analysis on H3GA0044562. (C) represents the 115 kb linkage disequilibrium block in the significant region on SSC15. (D) indicates the genotype effect plot of top SNP (H3GA0044562) related to SLOIN in 526 DLY pigs (** $P < 0.01$, ns $P > 0.05$).

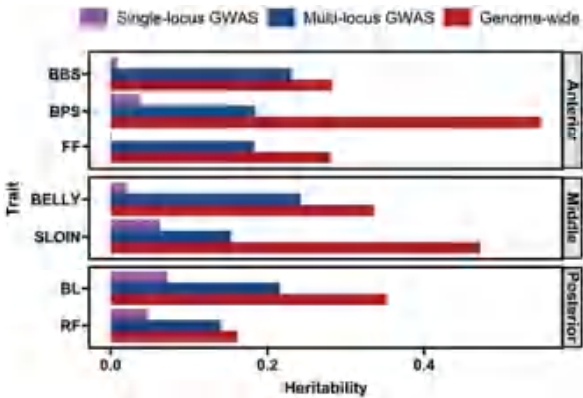


Figure 5. Heritability of seven carcass traits estimated from different SNP categories. The x-axis represents heritability, and the y-axis represents the anterior, middle, and posterior carcass traits. The anterior portion includes BBS (boneless butt shoulder), BPS (boneless picnic shoulder), and FF (front fat). The middle portion includes BELLY (belly) and SLOIN (skin-on whole loin). The posterior portion includes BL (boneless leg) and RF (rear fat). The red, purple, and blue colors indicate heritability estimated from the genome-wide SNPs, as well as the significant SNPs identified through single- and multi-locus GWAS, respectively.

Supplementary Tables S2 and S3). These findings demonstrate that multi-locus GWAS can effectively capture more genetic signals than single-locus GWAS (Wang et al., 2016; Wen et al., 2017; Ren et al., 2018; Tamba and Zhang, 2018; Zhang et al., 2020). Therefore, to overcome the deficiency of missing heritability, we performed a combination of single- and multi-locus models to improve the efficiency of the study. This strategy is consistent with previous studies; for example, Ding et al., (2019) used both the single-locus analysis and multi-locus analyses (mrMLM, FASTmrEMMA, and ISIS EM-BLASSO) to overcome some limitations of single-locus GWAS. Their study identified 10 candidate genes that appeared to have biochemical and physiological effects related to intramuscular fat in pigs. Zhuang et al., (2020) found that the combination of single- and multi-locus GWAS detected additional SNPs compared to only one model. Our study identified 177 nonredundant SNPs associated with seven carcass traits in 526 DLY pigs, 96% of which were determined by multi-locus GWAS. In tandem with these findings, many candidate genes were annotated using a series of bioinformatics analyses and functional annotations.

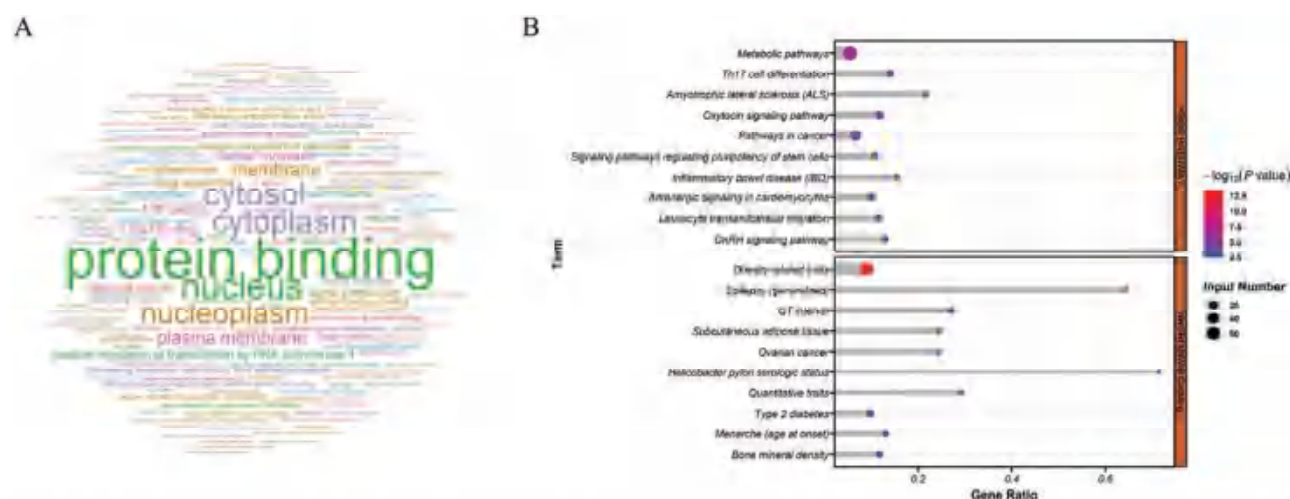


Figure 6. Plots of significant terms in GO, KEGG PATHWAY, and NHGRI GWAS Catalog. (A) represents word cloud of the GO enrichment of the biological process for protein-coding genes within a 1 Mb region centered on the GWAS signals (Corrected P -value < 0.01). (B) denotes the top 10 clusters of KEGG PATHWAY (top) and NHGRI GWAS catalog (bottom) enriched terms.

Genetics loci and candidate genes for carcass traits in anterior portion

BBS and BPS are the most consumed and preferable cuts in many countries (Choe et al., 2015; Hoa et al., 2021), both of which are obtained from the pig's shoulder (Figure 1A and Supplementary Table S1; Newcom et al., 2002). Our study has identified two novel genes associated with BBS. The *PPP3CA* gene encodes calcineurin, a widely distributed phosphatase that has a role in various physiological pathways, including the oxidative metabolism of skeletal muscle (Davoli et al., 2016). Differences in the structural and metabolic characteristics of skeletal muscle fibers tend to determine the meat transformation events in the muscle (da Costa et al., 2007). In addition, the *CPEB4* gene has been described to be associated with the human waist-hip ratio (Heid et al., 2010) and rib eye area in cattle (Mudadu et al., 2016). BPS is situated directly below the BBS, and the *ECH1* gene is regarded as the candidate gene that influences BPS. This gene encodes an enzyme that hydrates short- and medium-chain enoyl-CoA and is linked to the upregulation of β -oxidation (Bahnson et al., 2002). In general, indigenous Chinese pigs exhibit lower gene expression, resulting in better meat quality parameters than commercial pigs (Fernández-Barroso et al., 2022). Therefore, *PPP3CA*, *CPEB4*, and *ECH1* are important for meat quality and could be considered interesting candidate genes for BBS and BPS, respectively.

Fat content has a great impact on the texture and flavor of meat products. Our study also found two candidates (*CACNB2* and *ZNF217*) related to FF. The *CACNB2* gene was found to be associated with loin marbling score, lean cuts, and ham weight in a genome-wide analysis of Duroc pigs (Bertolini et al., 2018). Another candidate gene, *ZNF217*, has been shown to be associated with increased carcass weight in cows (Mullen et al., 2011), although its precise role in fat content is still not well-understood.

Genetics loci and candidate genes for carcass traits in middle portion

Pork belly is the most economically preferable among primal pork cuts due to the persistence of Asian cultural preference (Marcoux et al., 2007; Lee et al., 2018a). Fresh pork belly

offers high-quality protein from lean cuts and substantial micronutrients, including fat-soluble vitamins and minerals (Choe et al., 2015). We found that the *FGFRL1* gene downstream of WU_10.2_8_28850 (SSC8: 233,563 bp) may affect BELLY. *FGFRL1* plays an important role in the embryonic development of slow muscle fibers (Amann et al., 2014). Moreover, a previous study has reported that *FGFRL1* is associated with overgrowth (Matoso et al., 2014).

Pork loins with more marbling have been shown to be associated with greater juiciness, flavor, and tenderness scores (Fernandez et al., 1999; Liu et al., 2018). In this study, we identified one QTL for SLOIN. Previous researchers discovered several QTLs associated with loin muscle overlap with this QTL (Wimmers et al., 2007; Edwards et al., 2008; Choi et al., 2010), suggesting that confident candidate genes exist near this region. Consequently, we speculated that the *MYO3B* gene might be associated with SLOIN weight. This gene encodes one of the class III myosins (Merritt et al., 2012), and a recent study reported that variants in the *MYO3B* gene are involved in the visceral-to-subcutaneous fat ratio (Kim et al., 2022).

Genetics loci and candidate genes for carcass traits in posterior portion

BL is commonly used in ham production. For example, Xuanwei ham, Jinhua ham, and mutton ham are popular traditional dry-cured meat products in China (Wang et al., 2021b). Previous studies have identified candidate genes associated with dry-cured ham production traits, including *CACNB2* in Italian Duroc heavy pigs (Bertolini et al., 2018). Palombo et al. performed single-step GWAS for untrimmed and trimmed thigh weight in Italian crossbred pigs and discovered several candidate genes (*KCNA3*, *PROK1*, *CSF1*, and *AMPD2*; Palombo et al., 2021). In our study, ALGA0033122 (SSC5: 79,625,050 bp) associated with BL is located near the *CHST11* gene, which has been suggested to be involved in limb bud elongation and bone formation (Cortes et al., 2009).

For RF, we highlighted the *LRRK2* gene, which was previously recognized as an essential gene for lipid accumulation (Boddu et al., 2015) and a candidate gene associated with intramuscular fat content in Suhuai pigs (Wang et al., 2021a).

This gene was also identified as necessary for plasma with very low-density lipoprotein concentration using GWAS in chicken (Zhang et al., 2019).

Conclusion

This study performed a combined GWAS approach to investigate the weights of seven carcass components in commercial crossbred pigs. Our findings show that the integration of single- and multi-locus GWAS identified more large-effect SNPs in contrast to using only one model. Consequently, 24, 30, 26, 25, 17, 31, and 27 SNPs were found to be associated with BBS, BPS, FF, BELLY, SLOIN, BL, and RF, respectively. Furthermore, we successfully identified several candidate genes. In conclusion, the utilization of this combined GWAS approach revealed a comprehensive understanding of carcass traits and provided unique insights for further molecular breeding of pigs, particularly in the context of genomic selection.

Supplementary Data

Supplementary data are available at *Journal of Animal Science* online.

Acknowledgments

We would like to thank all staff at the pig core breeding farms of Wens Foodstuff Group Co., Ltd (Guangdong, China) for their help in sample collection. This research was supported by the Natural Science Foundation of Guangdong Province project (2018B030313011) and the Guangdong Modern Agricultural Industrial Technology System pig innovation team project (2022KJ26). The funders had no role in the study design, data collection, analysis, publication decision, or manuscript preparation.

Conflict of Interest Statement

The authors declare no real or perceived conflicts of interest.

Availability of Data and Material

The SNP genotyping data containing variant information for the DLY ($N = 526$) pigs are not publicly available because the genotyped animals belong to commercial breeding companies, but they can be obtained from the corresponding author under reasonable requests.

Authors' Contributions

J.Y. and E.Z. conceived and designed the experiment. Y.Q., Z.Z., F.M., D.R., C.X., F.M., L.P., R.D., G.C., E.Z., and M.Y. collected the samples, recorded the phenotypes, and performed the experiments. Y.Q., Z.Z., and F.M. analyzed the data. Y.Q., Z.Z., J.Y., and E.Z. wrote the manuscript. Z.W. contributed to the materials. All authors reviewed and approved the manuscript.

Literature Cited

Alfaia, C. M., P. A. Lopes, M. S. Madeira, J. M. Pestana, D. Coelho, F. Toldrá, and J. A. M. Prates. 2019. Chapter Two - Current feeding strategies to improve pork intramuscular fat content and its nutritional quality. In: F. Toldrá, editor. *Advances in Food and Nutrition*



- Research*. Vol. 89. Academic Press. p. 53–94. Available from: <https://www.sciencedirect.com/science/article/pii/S104345261930035X>
- Amann, R., S. Wyder, A. M. Slavotinek, and B. Trueb. 2014. The FgfrL1 receptor is required for development of slow muscle fibers. *Dev. Biol.* 394:228–241. doi:10.1016/j.ydbio.2014.08.016
- Bahnson, B. J., V. E. Anderson, and G. A. Petsko. 2002. Structural mechanism of enoyl-CoA hydratase: three atoms from a single water are added in either an E1cb stepwise or concerted fashion. *Biochemistry* 41:2621–2629. doi:10.1021/bi015844p
- Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265. doi:10.1093/bioinformatics/bth457
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol.* 57:289–300. doi:10.1111/j.2517-6161.1995.tb02031.x
- Bertolini, F., G. Schiavo, G. Galimberti, S. Bovo, M. D'Andrea, M. Gallo, L. Buttazzoni, M. F. Rothschild, and L. Fontanesi. 2018. Genome-wide association studies for seven production traits highlight genomic regions useful to dissect dry-cured ham quality and production traits in Duroc heavy pigs. *Animal*. 12:1777–1784. doi:10.1017/S1751731118000757
- Boddu, R., T. D. Hull, S. Bolisetty, X. Hu, M. S. Moehle, J. P. L. Daher, A. I. Kamal, R. Joseph, J. F. George, A. Agarwal, et al. 2015. Leucine-rich repeat kinase 2 deficiency is protective in rhabdomyolysis-induced kidney injury. *Hum. Mol. Genet.* 24:4078–4093. doi:10.1093/hmg/ddv147
- Bu, D., H. Luo, P. Huo, Z. Wang, S. Zhang, Z. He, Y. Wu, L. Zhao, J. Liu, J. Guo, S. Fang, W. Cao, L. Yi, Y. Zhao, and L. Kong. 2021. KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Research*. 49:W317–W325. doi:10.1093/nar/gkab447
- Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaSci.* 4:7. doi:10.1186/s13742-015-0047-8
- Chen, G., Y. Cai, Y. Su, D. Wang, X. Pan, and X. Zhi. 2021. Study of meat quality and flavour in different cuts of Duroc-Bamei binary hybrid pigs. *Vet. Med. Sci.* 7:724–734. doi:10.1002/vms3.409
- Cho, I. -C., H. -B. Park, J. S. Ahn, S. -H. Han, J. -B. Lee, H. -T. Lim, C. -K. Yoo, E. -J. Jung, D. -H. Kim, W. -S. Sun, et al. 2019. A functional regulatory variant of MYH3 influences muscle fiber-type composition and intramuscular fat content in pigs. *PLoS Genet.* 15:e1008279. doi:10.1371/journal.pgen.1008279
- Choe, J. -H., H. -S. Yang, S. -H. Lee, and G. -W. Go. 2015. Characteristics of pork belly consumption in South Korea and their health implication. *J Anim Sci Technol.* 57:22. doi:10.1186/s40781-015-0057-1
- Choi, I., J. P. Steibel, R. O. Bates, N. E. Raney, J. M. Rumph, and C. W. Ernst. 2010. Application of alternative models to identify QTL for growth traits in an F2 Duroc x Pietrain pig resource population. *BMC Genet.* 11:1–10. doi:10.1186/1471-2156-11-97
- Cortes, M., A. T. Baria, and N. B. Schwartz. 2009. Sulfation of chondroitin sulfate proteoglycans is necessary for proper Indian hedgehog signaling in the developing growth plate. *Development*. 136:1697–1706. doi:10.1242/dev.030742
- da Costa, N., J. Edgar, P. -T. Ooi, Y. Su, J. D. Meissner, and K. -C. Chang. 2007. Calcineurin differentially regulates fast myosin heavy chain genes in oxidative muscle fibre type conversion. *Cell Tissue Res.* 329:515–527. doi:10.1007/s00441-007-0441-3
- Davoli, R., D. Luise, V. Mingazzini, P. Zambonelli, S. Braglia, A. Serra, and V. Russo. 2016. Genome-wide study on intramuscular fat in Italian Large White pig breed using the PorcineSNP60 BeadChip. *J. Anim. Breed. Genet.* 133:277–282. doi:10.1111/jbg.12189
- Deary, I. J., J. Yang, G. Davies, S. E. Harris, A. Tenesa, D. Liewald, M. Luciano, L. M. Lopez, A. J. Gow, J. Corley, et al. 2012. Genetic contributions to stability and change in intelligence from childhood to old age. *Nature* 482:212–215. doi:10.1038/nature10781
- Devlin, B., and K. Roeder. 1999. Genomic control for association studies. *Biometrics* 55:997–1004. doi:10.1111/j.0006-341x.1999.00997.x

- Ding, R., M. Yang, J. Quan, S. Li, Z. Zhuang, S. Zhou, E. Zheng, L. Hong, Z. Li, G. Cai, et al. 2019. Single-Locus and Multi-Locus Genome-Wide Association Studies for Intramuscular Fat in Duroc Pigs. *Front. Genet.* 10:619. doi:10.3389/fgene.2019.00619
- Ding, R., Y. Qiu, Z. Zhuang, D. Ruan, J. Wu, S. Zhou, J. Ye, L. Cao, L. Hong, Z. Xu, et al. 2021. Genome-wide association studies reveals polygenic genetic architecture of litter traits in Duroc pigs. *Theriogenology*. 173:269–278. doi:10.1016/j.theriogenology.2021.08.012
- Edwards, D. B., C. W. Ernst, N. E. Raney, M. E. Doumit, M. D. Hoge, and R. O. Bates. 2008. Quantitative trait locus mapping in an F2 Duroc × Pietrain resource population: II. Carcass and meat quality traits1. *J. Anim. Sci.* 86:254–266. doi:10.2527/jas.2006-626
- Fernandez, X., G. Monin, A. Talmant, J. Mourot, and B. Lebret. 1999. Influence of intramuscular fat content on the quality of pig meat - 1. Composition of the lipid fraction and sensory characteristics of m. longissimus lumborum. *Meat Sci.* 53:59–65. doi:10.1016/s0309-1740(99)00037-6
- Fernández-Barroso, M., J. M. García-Casco, Y. Núñez, L. Ramírez-Hidalgo, G. Matos, and M. Muñoz. 2022. Understanding the role of myoglobin content in Iberian pigs fattened in an extensive system through analysis of the transcriptome profile. *Anim. Genet.* 53:352–367. doi:10.1111/age.13195
- Gabriel, S. B., S. F. Schaffner, H. Nguyen, J. M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, M. Faggart, et al. 2002. The structure of haplotype blocks in the human genome. *Science* 296:2225–2229. doi:10.1126/science.1069424
- Galvan, A., J. P. A. Ioannidis, and T. A. Draganí. 2010. Beyond genome-wide association studies: genetic heterogeneity and individual predisposition to cancer. *Trends Genet.* 26:132–141. doi:10.1016/j.tig.2009.12.008
- Heid, I. M., A. U. Jackson, J. C. Randall, T. W. Winkler, L. Qi, V. Steinthorsdottir, G. Thorleifsson, M. C. Zillikens, E. K. Speliotes, R. Mägi, et al. 2010. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat. Genet.* 42:949–960. doi:10.1038/ng.685
- Hoa, V. -B., K. -H. Seol, S. -M. Kang, Y. -S. Kim, and S. -H. Cho. 2021. A study on shelf life of prepackaged retail-ready Korean native black pork belly and shoulder butt slices during refrigerated display. *Anim. Biosci.* 34:2012–2022. doi:10.5713/ab.21.0172
- Kim, H. -J., H. Son, J. Sung, J. M. Yun, H. Kwon, B. Cho, J. -I. Kim, and J. -H. Park. 2022. A genome-wide association study on abdominal adiposity-related traits in adult Korean men. *Obes. Facts.* 15:590–599. doi:10.1159/000524670
- Lee, S. H., J. Yang, M. E. Goddard, P. M. Visscher, and N. R. Wray. 2012. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics.* 28:2540–2542. doi:10.1093/bioinformatics/bts474
- Lee, E. -A., J. -H. Kang, J. -H. Cheong, K. -C. Koh, W. -M. Jeon, J. -H. Choe, K. -C. Hong, and J. -M. Kim. 2018a. Evaluation of whole pork belly qualitative and quantitative properties using selective belly muscle parameters. *Meat Sci.* 137:92–97. doi:10.1016/j.meatsci.2017.11.012
- Lee, S. H., M. H. Hur, E. A. Lee, K. C. Hong, and J. M. Kim. 2018b. Genomic characterization of the porcine CRT3 and the effects of a non-synonymous mutation p.V515F on lean meat production and belly fat. *Meat Sci.* 137:211–215. doi:10.1016/j.meatsci.2017.11.019
- Li, C., Y. Fu, R. Sun, Y. Wang, and Q. Wang. 2018. Single-locus and multi-locus genome-wide association studies in the genetic dissection of fiber quality traits in upland cotton (*Gossypium hirsutum* L.). *Front. Plant Sci.* 9:1083. doi:10.3389/fpls.2018.01083
- Li, J., S. Peng, L. Zhong, L. Zhou, G. Yan, S. Xiao, J. Ma, and L. Huang. 2021. Identification and validation of a regulatory mutation upstream of the BMP2 gene associated with carcass length in pigs. *Genet. Sel. Evol.* 53:94. doi:10.1186/s12711-021-00689-0
- Liu, J. -H., X. Sun, J. M. Young, L. A. Bachmeier, and D. J. Newman. 2018. Predicting pork loin intramuscular fat using computer vision system. *Meat Sci.* 143:18–23. doi:10.1016/j.meatsci.2018.03.020
- Ma, J., J. Yang, L. Zhou, J. Ren, X. Liu, H. Zhang, B. Yang, Z. Zhang, H. Ma, X. Xie, 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:e1004710. doi:10.1371/journal.pgen.1004710
- Maher, B. 2008. Personal genomes: the case of the missing heritability. *Nature* 456:18–21. doi:10.1038/456018a
- Marcoux, M., C. Pomar, L. Faucitano, and C. Brodeur. 2007. The relationship between different pork carcass lean yield definitions and the market carcass value. *Meat Sci.* 75:94–102. doi:10.1016/j.meatsci.2006.07.001
- Marsaglia, G., W. W. Tsang, and J. Wang. 2003. Evaluating kolmogorov's distribution. *J. Stat. Soft.* 8. doi:10.18637/jss.v008.i18. Available from: <http://www.jstatsoft.org/v08/i18/>
- Matoso, E., F. Ramos, J. Ferrão, L. M. Pires, A. Mascarenhas, J. B. Melo, and I. M. Carreira. 2014. Interstitial 287 kb deletion of 4p16.3 including FGFR1 gene associated with language impairment and overgrowth. *Mol. Cytogenet.* 7:87. doi:10.1186/s13039-014-0087-2
- Merritt, R. C., U. Manor, F. T. Salles, M. Grati, A. C. Dose, W. C. Unrath, O. A. Quintero, C. M. Yengo, and B. Kachar. 2012. Myosin IIIB uses an actin-binding motif in its espin-1 cargo to reach the tips of actin protrusions. *Curr. Biol.* 22:320–325. doi:10.1016/j.cub.2011.12.053
- Mikawa, S., S. Sato, M. Nii, T. Morozumi, G. Yoshioka, N. Imaeda, T. Yamaguchi, T. Hayashi, and T. Awata. 2011. Identification of a second gene associated with variation in vertebral number in domestic pigs. *BMC Genet.* 12:51–13. doi:10.1186/1471-2156-12-5
- Mudadu, M. A., L. R. Porto-Neto, F. B. Mokry, P. C. Tizioto, P. S. N. Oliveira, R. R. Tullio, R. T. Nassu, S. C. M. Niciura, P. Tholon, M. M. Alencar, et al. 2016. Genomic structure and marker-derived gene networks for growth and meat quality traits of Brazilian Nelore beef cattle. *BMC Genomics* 17:235. doi:10.1186/s12864-016-2535-3
- Mullen, M., D. Berry, D. Howard, M. Diskin, C. Lynch, L. Giblin, D. Kenny, D. Magee, K. Meade, and S. Waters. 2011. Single Nucleotide Polymorphisms in the Insulin-Like Growth Factor 1 (IGF-1) Gene are Associated with Performance in Holstein-Friesian Dairy Cattle. *Front. Genet.* 2:3. doi:10.3389/fgene.2011.00003
- Newcom, D. W., T. J. Baas, J. W. Mabry, and R. N. Goodwin. 2002. Genetic parameters for pork carcass components. *J. Anim. Sci.* 80:3099–3106. doi:10.2527/2002.80123099x
- Nosková, A., M. Hiltbold, F. Janett, T. Echtermann, Z. -H. Fang, X. Siller, C. Selige, A. Hofer, S. Neuenschwander, and H. Pausch. 2020. Infertility due to defective sperm flagella caused by an intronic deletion in DNAH17 that perturbs splicing. *Genetics* 217:iyaa033. doi:10.1093/genetics/iyaa033
- Palombo, V., M. D'Andrea, D. Licastro, S. Dal Monego, S. Sgorlon, M. Sandri, and B. Stefanon. 2021. Single-step genome wide association study identifies QTL signals for untrimmed and trimmed thigh weight in Italian crossbred pigs for dry-cured ham production. *Animals*. 11:1612. doi:10.3390/ani11061612
- Price, A. L., N. A. Zaitlen, D. Reich, and N. Patterson. 2010. New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* 11:459–463. doi:10.1038/nrg2813
- Pritchard, J. K., and N. A. Rosenberg. 1999. Use of Unlinked Genetic Markers to Detect Population Stratification in Association Studies. *Am. J. Hum. Genet.* 65:220–228. doi:10.1086/302449
- Qiu, Y., R. Ding, Z. Zhuang, J. Wu, M. Yang, S. Zhou, Y. Ye, Q. Geng, Z. Xu, S. Huang, et al. 2021. Genome-wide detection of CNV regions and their potential association with growth and fatness traits in Duroc pigs. *BMC Genomics*. 22:332. doi:10.1186/s12864-021-07654-7
- Reich, D. E., and D. B. Goldstein. 2001. Detecting association in a case-control study while correcting for population stratification. *Genet. Epidemiol.* 20:4–16. doi:10.1002/1098-2272(200101)20:1<4::AID-GEPI2>3.0.CO;2-T
- Ren, W. -L., Y. -J. Wen, J. M. Dunwell, and Y. -M. Zhang. 2018. pK-WmEB: integration of Kruskal-Wallis test with empirical Bayes under polygenic background control for multi-locus genome-wide

- association study. *Heredity* 120:208–218. doi:[10.1038/s41437-017-0007-4](https://doi.org/10.1038/s41437-017-0007-4)
- Rohrer, G. A., and J. W. Keele. 1998. Identification of quantitative trait loci affecting carcass composition in swine: II. Muscling and wholesale product yield traits. *J. Anim. Sci.* 76:2255–2262. doi:[10.2527/1998.7692255x](https://doi.org/10.2527/1998.7692255x)
- Segura, V., B. J. Vilhjálmsson, A. Platt, A. Korte, U. Seren, Q. Long, and M. Nordborg. 2012. An efficient multi-locus mixed model approach for genome-wide association studies in structured populations. *Nat. Genet.* 44:825–830. doi:[10.1038/ng.2314](https://doi.org/10.1038/ng.2314)
- Tamba, C. L., and Y.-M. Zhang. 2018. A fast mrMLM algorithm for multi-locus genome-wide association studies. *bioRxiv*, doi:[10.1101/341784](https://doi.org/10.1101/341784), preprint: not peer reviewed.
- Wang, S. -B., J. -Y. Feng, W. -L. Ren, B. Huang, L. Zhou, Y. -J. Wen, J. Zhang, J. M. Dunwell, S. Xu, and Y. -M. Zhang. 2016. Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. *Sci. Rep.* 6:19444. doi:[10.1038/srep19444](https://doi.org/10.1038/srep19444)
- Wang, B. B., L. M. Hou, W. D. Zhou, H. Liu, W. Tao, W. J. Wu, P. P. Niu, Z. P. Zhang, J. Zhou, Q. Li, et al. 2021a. Genome-wide association study reveals a quantitative trait locus and two candidate genes on Sus scrofa chromosome 5 affecting intramuscular fat content in Suhuai pigs. *Animal*. 15:100341. doi:[10.1016/j.animal.2021.100341](https://doi.org/10.1016/j.animal.2021.100341)
- Wang, J., M. Guo, Q. Wang, J. Dong, S. Lu, B. Lyu, and X. Ma. 2021b. Antioxidant activities of peptides derived from mutton ham, Xuanwei ham and Jinhua ham. *Food Res. Int.* 142:110195. doi:[10.1016/j.foodres.2021.110195](https://doi.org/10.1016/j.foodres.2021.110195)
- Wen, Y. -J., H. Zhang, Y. -L. Ni, B. Huang, J. Zhang, J. -Y. Feng, S. -B. Wang, J. M. Dunwell, Y. -M. Zhang, and R. Wu. 2017. Methodological implementation of mixed linear models in multi-locus genome-wide association studies. *Brief. Bioinform.* 19:700–712. doi:[10.1093/bib/bbw145](https://doi.org/10.1093/bib/bbw145)
- Wimmers, K., E. Murani, M. F. W. Te Pas, K. C. Chang, R. Davoli, J. W. M. Merks, H. Henne, M. Muraniova, N. Da Costa, B. Harlizius, et al. 2007. Associations of functional candidate genes derived from gene-expression profiles of prenatal porcine muscle tissue with meat quality and muscle deposition. *Animal Genetics*. 38:474–484. doi:[10.1111/j.1365-2052.2007.01639.x](https://doi.org/10.1111/j.1365-2052.2007.01639.x)
- Wu, P., K. Wang, J. Zhou, D. Chen, A. Jiang, Y. Jiang, L. Zhu, X. Qiu, X. Li, and G. Tang. 2021. A combined GWAS approach reveals key loci for socially-affected traits in Yorkshire pigs. *Commun. Biol.* 4:1–11. doi:[10.1038/s42003-021-02416-3](https://doi.org/10.1038/s42003-021-02416-3)
- Yang, Q., J. Cui, I. Chazaro, L. A. Cupples, and S. Demissie. 2005. Power and type I error rate of false discovery rate approaches in genome-wide association studies. *BMC Genet.* 6:S134. doi:[10.1186/1471-2156-6-S1-S134](https://doi.org/10.1186/1471-2156-6-S1-S134)
- Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher. 2011. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88:76–82. doi:[10.1016/j.ajhg.2010.11.011](https://doi.org/10.1016/j.ajhg.2010.11.011)
- Yang, J., T. Ferreira, A. P. Morris, S. E. Medland, P. A. F. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. N. Weedon, R. J. Loos, et al. 2012. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* 44:369–3S3. doi:[10.1038/ng.2213](https://doi.org/10.1038/ng.2213)
- Yang, J., L. Huang, M. Yang, Y. Fan, L. Li, S. Fang, W. Deng, L. Cui, Z. Zhang, H. Ai, et al. 2016. Possible introgression of the VRTN mutation increasing vertebral number, carcass length and teat number from Chinese pigs into European pigs. *Sci. Rep.* 6:19240. doi:[10.1038/srep19240](https://doi.org/10.1038/srep19240)
- Zhang, H., L. Shen, Y. Li, Z. Xu, X. Zhang, J. Yu, Z. Cao, and P. Luan. 2019. Genome-wide association study for plasma very low-density lipoprotein concentration in chicken. *J. Anim. Breed. Genet.* 136:351–361. doi:[10.1111/jbg.12397](https://doi.org/10.1111/jbg.12397)
- Zhang, Y. -W., C. L. Tamba, Y. -J. Wen, P. Li, W. -L. Ren, Y. -L. Ni, J. Gao, and Y. -M. Zhang. 2020. mrMLM v4.0.2: An R Platform for Multi-locus Genome-wide Association Studies. *Genomics Proteomics Bioinformatics*. 18:481–487. doi:[10.1016/j.gpb.2020.06.006](https://doi.org/10.1016/j.gpb.2020.06.006)
- Zhang, H., Z. Zhuang, M. Yang, R. Ding, J. Quan, S. Zhou, T. Gu, Z. Xu, E. Zheng, G. Cai, et al. 2021. Genome-wide detection of genetic loci and candidate genes for body conformation traits in duroc × landrace × yorkshire crossbred pigs. *Front. Genet.* 12:664343. doi:[10.3389/fgene.2021.664343](https://doi.org/10.3389/fgene.2021.664343)
- Zhou, X., and M. Stephens. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44:821–824. doi:[10.1038/ng.2310](https://doi.org/10.1038/ng.2310)
- Zhou, S., R. Ding, F. Meng, X. Wang, Z. Zhuang, J. Quan, Q. Geng, J. Wu, E. Zheng, Z. Wu, et al. 2021. A meta-analysis of genome-wide association studies for average daily gain and lean meat percentage in two Duroc pig populations. *BMC Genomics* 22:12. doi:[10.1186/s12864-020-07288-1](https://doi.org/10.1186/s12864-020-07288-1)
- Zhuang, Z., R. Ding, L. Peng, J. Wu, Y. Ye, S. Zhou, X. Wang, J. Quan, E. Zheng, G. Cai, et al. 2020. Genome-wide association analyses identify known and novel loci for teat number in Duroc pigs using single-locus and multi-locus models. *BMC Genomics* 21:344. doi:[10.1186/s12864-020-6742-6](https://doi.org/10.1186/s12864-020-6742-6)

Article

Integrated Single-Trait and Multi-Trait GWASs Reveal the Genetic Architecture of Internal Organ Weight in Pigs

Xuehua Li ^{1,†} , Jie Wu ^{1,†}, Zhanwei Zhuang ¹, Yong Ye ¹, Shenping Zhou ¹, Yibin Qiu ¹, Donglin Ruan ¹, Shiyuan Wang ¹, Jie Yang ^{1,2} , Zhenfang Wu ^{1,2,3}, Gengyuan Cai ^{1,2,*} and Enqin Zheng ^{1,2,*}

¹ College of Animal Science and National Engineering Research Center for Breeding Swine Industry, South China Agricultural University, Guangzhou 510642, China

² Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, South China Agricultural University, Guangzhou 510642, China

³ Yunfu Subcenter of Guangdong Laboratory for Lingnan Modern Agriculture, Yunfu 527400, China

* Correspondence: cgy0415@163.com (G.C.); eqzheng@scau.edu.cn (E.Z.)

† These authors contributed equally to this work.

Simple Summary: In this study, bioinformatics approaches were used to better understand the genetic architecture of internal organ weights in three-way crossbred commercial pigs and to map genetic markers and genes. For this purpose, we used single-trait and multi-trait genome-wide association studies (GWASs) followed by a haplotype block analysis. We explored the key genetic markers and genes from the internal organ weight genome-wide association study results of three-way crossbred commercial pigs. In this manner, five genes, *TPK1*, *POU6F2*, *PBX3*, *UNC5C*, and *BMPR1B*, were defined as central in affecting internal organ weight in pigs. Moreover, *APK1*, *ANO6*, and *UNC5C* were identified to be pleiotropic in multi-trait GWASs. These results can be applied to various types of genomic studies of pigs.

Abstract: Internal organ weight is an essential indicator of growth status as it reflects the level of growth and development in pigs. However, the associated genetic architecture has not been well explored because phenotypes are difficult to obtain. Herein, we performed single-trait and multi-trait genome-wide association studies (GWASs) to map the genetic markers and genes associated with six internal organ weight traits (including heart weight, liver weight, spleen weight, lung weight, kidney weight, and stomach weight) in 1518 three-way crossbred commercial pigs. In summation, single-trait GWASs identified a total of 24 significant single-nucleotide polymorphisms (SNPs) and 5 promising candidate genes, namely, *TPK1*, *POU6F2*, *PBX3*, *UNC5C*, and *BMPR1B*, as being associated with the six internal organ weight traits analyzed. Multi-trait GWAS identified four SNPs with polymorphisms localized on the *APK1*, *ANO6*, and *UNC5C* genes and improved the statistical efficacy of single-trait GWASs. Furthermore, our study was the first to use GWASs to identify SNPs associated with stomach weight in pigs. In conclusion, our exploration of the genetic architecture of internal organ weights helps us better understand growth traits, and the key SNPs identified could play a potential role in animal breeding programs.

Keywords: internal organ weight; GWAS; DLY pigs; genetic architecture



Citation: Li, X.; Wu, J.; Zhuang, Z.; Ye, Y.; Zhou, S.; Qiu, Y.; Ruan, D.; Wang, S.; Yang, J.; Wu, Z.; et al. Integrated Single-Trait and Multi-Trait GWASs Reveal the Genetic Architecture of Internal Organ Weight in Pigs. *Animals* **2023**, *13*, 808. <https://doi.org/10.3390/ani13050808>

Academic Editor: Krzysztof Flisikowski

Received: 7 January 2023

Revised: 18 February 2023

Accepted: 20 February 2023

Published: 23 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Body weight, which can reflect growth performance and thus affect economic efficiency, has attracted a lot of attention in animal breeding programs. The body weight of cattle is the sum of various elements, including fat weight, internal organ weight, muscle weight, and bone weight, among others. Of these components, internal organ weight constitutes 14% of the total body weight of cattle [1]. The weight and size of an organ are salient features that serve as dependable predictors of its developmental progression, wherein an augmented organ mass typically alludes to a heightened degree of maturation.

Accelerated organ development leads to a smoother coordination of internal organs during vital biological processes such as oxygen transport, blood circulation, lipid metabolism, and digestion. This refinement of these processes can positively impact growth and economic traits. Previous studies have shown that the internal organ weights of crossbred steer calves are strongly correlated with carcass growth rate [2]. Moreover, in humans, internal organ weights have been shown to be positively correlated with body weight and height in normal Zambian adults [3]. Thus, comprehending the genetic architecture of heart weight (Heart WT), liver weight (Liver WT), spleen weight (Spleen WT), lung weight (Lung WT), kidney weight (Kidney WT), and stomach weight (Stomach WT) will propel genetic progress and facilitate the successful implementation of breeding programs.

Genome-wide association studies (GWASs) are widely used to identify quantitative trait loci (QTL) and candidate genes associated with complex traits in animals and plants. To date, the number of QTL associated with Heart WT, Liver WT, Spleen WT, Lung WT, and Kidney WT are 29, 31, 19, 5 and 8, respectively, and no QTL have been reported to be associated with Stomach WT in the pig QTL database [4] (accessed on 15 November 2022). Previous studies reported 39 QTL to be associated with internal organ weight in four local pig populations and one commercial population [5]. For instance, Zhang et al. [6] showed that a 2 cM QTL on *Sus scrofa* chromosome 2 (SSC2) was significantly associated with Heart WT, and three QTL were associated with Liver WT, Lung WT, and Spleen WT. Although several studies have identified QTL to be associated with internal organ weight [7,8], the process of genetic improvement remains slow.

The difficulty (and high cost) of obtaining phenotypes for internal organ weight studies has led to fewer studies on its genetic architecture. Moreover, previous studies conducted single-trait GWASs for internal organ weight to map the genetic markers and genes; however, internal organ development is mutually coordinated by each different organ, and the single-nucleotide polymorphisms (SNPs) in the genome may act on multiple organs at the same time. Therefore, it is difficult to identify SNPs and candidate genes that affect multiple internal organs simultaneously using single-trait GWASs. Therefore, herein, we performed multi-trait GWASs to identify polymorphic SNPs and improve statistical efficiency, which mainly depends on the genetic correlation between traits [9,10]. In this manner, it was observed that the statistical efficiency was improved in the case of low trait correlations [11,12].

Previous studies demonstrated the superiority of conducting multi-trait GWASs in terms of uncovering the genetic architecture of complex traits in animals. For instance, Zhou et al. [13] performed multi-trait GWASs to identify 21 pleiotropic SNPs that were not detected via single-trait GWASs in three body size traits. In Simmental beef cattle, An et al. [14] detected 29 pleiotropic SNPs that were functional in all three growth periods using multi-trait GWASs. To date, there are no studies that use multi-trait GWASs to analyze the genetic architecture of visceral weight. Herein, we performed multi-trait GWASs to compensate for the deficiencies associated with single-trait GWASs and to provide new insights into the genetic mechanisms of multi-organ co-development.

The aim of this study was to map the genetic markers and candidate genes associated with internal organ weight in pigs. To this end, we conducted single-trait and multi-trait GWASs for six internal organ weight traits in 1518 crossbred commercial Duroc × (Landrace × Yorkshire) DLY pigs. The results from the current study advanced our understanding of the genetic basis for internal organ weight and further revealed the complexity of the genetic architecture of internal organ weight in pigs. Integrating SNP results from GWASs as a source of prior biological information in the improvement program enhances the selection process by assigning higher weight to key SNPs that are critical for improving internal organ weight traits.

2. Materials and Methods

2.1. Ethical Statement

All animals used in this study were treated in accordance with the guidelines for the use of laboratory animals of the Ministry of Agriculture of China and with the approval of South China Agricultural University (Guangzhou, China), No. 2018F089.

2.2. Animal Samples and Phenotype Collection

Experimental animals were selected from a DLY three-way crossbred commercial line with no overlapping blood relations, through random selection based on genealogy, in which 89 Duroc boars were mated with 397 Landrace \times Yorkshire sows to produce a large number of offspring. All pigs were raised in four farms of the Guangdong Wens Food Group Co., Ltd. (Guangzhou, China). In brief, a total of 1518 individuals (757 boars and 764 sows) were reared with free access to water and feed and were fattened to 115 kg. They were euthanized in 13 batches with a 24 h interval between each batch and had an average slaughter age of about 7 months. After the pigs were euthanized, their phenotypes were recorded, and their internal organs were excised, emptied, flushed, blotted dry, and weighed immediately using an electronic scale with a range of 0.0 kg to 300 kg and accuracy of ± 100 g. The scale was calibrated using the linear calibration method with 20% MAX or 60% MAX weight. The organ distribution is shown in Figure 1. R 4.2.1 software was used to test the normal distribution of the descriptive statistics of the internal organ traits.

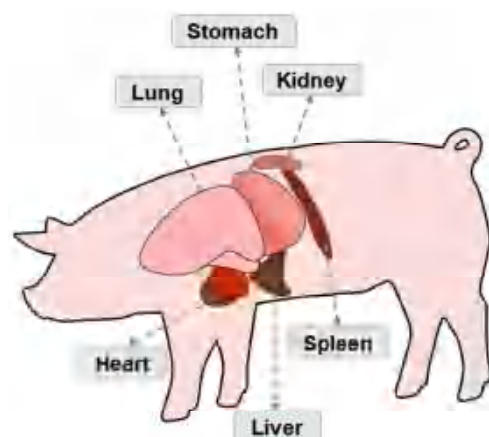


Figure 1. Distribution of heart, liver, spleen, lung, kidney, and stomach in the pig.

2.3. Genotyping and Quality Control

Ear samples were collected from all 1518 individuals, and genomic DNA was extracted from the ear tissue of each pig using a standard phenol–chloroform method and subsequently diluted to 50 ng/ μ L for the genotyping procedure, controlling the quality OD_{260/280} between 1.8 and 2.0. The 1518 DLY pigs were genotyped using the GeneSeek Porcine 50K SNP BeadChip (Neogen, Lincoln, NE, USA), which contained 50,703 SNPs. After genotyping, to ensure the accuracy and validity of the GWAS results, we performed a quality control (QC) procedure using the PLINK v1.07 software [15] with the following parameters: individual call rate > 95%; SNP call rate > 99%; minor allele frequency > 1%; and $p > 10^{-6}$ for the Hardy–Weinberg equilibrium test. Moreover, SNPs in sex chromosomes and unmapped regions were excluded. After QC, a final set of 31,941 eligible SNPs remained for subsequent single-trait and multi-trait GWASs.

2.4. Population Structure and Linkage Disequilibrium (LD) Estimation

PCA was conducted using the GCTA software [16] to assess the population structure, and PLINK v1.07 was used to calculate the LD decay distance, which was evaluated as the squared correlation of alleles (r^2) with a window size of 1000.

2.5. Single-Trait and Multi-Trait Genome-Wide Association Studies

The GEMMA software [17] was used to implement the linear mixed model (LMM) for the single-trait GWAS of each internal organ weight trait, including heart weight, liver weight, spleen weight, lung weight, kidney weight, and stomach weight. GEMMA calculated the genomic relatedness matrix (GRM) between individuals to account for the population structure. The mixed linear model was as follows:

$$y = W\alpha + X\beta + u + \varepsilon$$

where y is a vector of phenotypic values for each internal organ weight; W is the correlation matrix of covariates (fixed effects), including the top five eigenvectors of PCA, farm, sex, and slaughter lot; α is a vector of corresponding coefficients including the intercept; X is the genotypic vector of the SNP markers; β denotes the effect size of the SNP markers; u is a random effects vector, $u \sim MVNn(0, \lambda\tau^{-1}K)$; ε is the residual vector, $\varepsilon \sim MVNn(0, \tau^{-1}I_n)$; λ is the ratio of the specified variance components; τ^{-1} is the variance of the residuals; K denotes the kinship matrix; I is the unit matrix; n is the number of individuals in the DLY population; $MVNn$ denotes the multi-dimensional normal distribution.

Moreover, the GEMMA software [17] was used to implement the multivariate linear mixed models (mvLMMs) [18] for multi-trait GWASs to assess pleiotropic SNPs. The mvLMMs and LMMs were both implemented as described in previous studies [19]. In the current study, the LMMs and mvLMMs in the single-trait GWAS and the multi-trait GWAS utilized the same covariates. The multivariate linear mixed models were as follows:

$$Y = WA + x\beta^T + U + E; G \sim MN_{n \times d}(0, K, V_g), E \sim MN_{n \times d}(0, I_{n \times n}, V_e)$$

where Y is a matrix of six internal organs for 1518 individuals; W is a covariable matrix (fixed effects); A is a matrix of the corresponding coefficients; x is a vector that marks the genotypes; β is a vector of marker effect sizes for six internal organs' weights. U denotes the random effects; E is a matrix of errors; K denotes the kinship matrix; V_g denotes symmetric matrix of genetic variance component; I is an identity matrix; V_e denotes a symmetric matrix of the environmental variance component; $MN_{n \times d}(0, V_1, V_2)$ denotes the $n \times d$ matrix normal distribution with mean 0; V_1 denotes row covariance matrix; V_2 denotes column covariance matrix.

Furthermore, the Bonferroni correction can lead to an overcorrection and can be too conservative, this can result in a limited number of labeled association p -values that meet the standard across the genome. This can lead to a high false-negative rate. To address this issue, the false-discovery rate (FDR) was employed as a correction to the threshold [20]. Thus, the threshold p -value was calculated as $P = FDR * \frac{N}{M}$; the FDR was set to 0.01, N is the number of SNPs with p -value less than 0.01, and M refers to the total number of SNPs after quality control. Moreover, quantile–quantile (Q–Q) plots were constructed for the six internal organ weight traits to further assess the population structure.

In addition, the PLINK v1.07 and Haploview v4.2 software [21] were implemented to perform the haplotype block analysis in chromosomal regions with multiple significant SNPs. The default parameters of Haploview 4.2 [22] ($MAF > 0.05$, Mendelian error < 2 , and p -value $< 10^{-3}$ for the HWE test) were used to define the linkage disequilibrium (LD) blocks of SNPs.

2.6. Estimation of Heritability and Phenotypic Variation

In the present study, the restricted maximum likelihood (REML) method was used to assess the SNP-based heritability of each internal organ weight trait, and the percentage of phenotypic variation that could be explained by significant SNPs was calculated using GCTA software. SNP-based heritability and the percentage of phenotypic variation explained by significant SNPs were calculated as follows [23]:

$$y = X\beta + g + \varepsilon \text{ with } var(y) = A_g\sigma_g^2 + I\sigma_\varepsilon^2$$

where y is the phenotypic value of each internal organ weight trait; β is the vector of fixed effects, including the top five eigenvectors of PCA, farm, sex, and slaughter lot; X is an association matrix; g is the vector of total genetic effect of all the qualified SNPs for the 1518 DLY pigs; A_g is the genomic association matrix between different individuals; σ_g^2 is the additive genetic variance captured by either the genome-wide SNPs or the selected SNPs; σ_e^2 refers to residual variance.

2.7. Candidate Gene Search and Function Analysis

Our previous studies on this population showed that the average r^2 of 0.2 is about 200 kb apart [24]; the range for searching for the functional gene closest to the position of the significant SNP is determined based on the LD decay distance ($r^2 = 0.2$) of the populations [25]. We used the “biomaRt” package [26] in R, based on the Sus scrofa 11.1 genome version database (http://ensembl.org/Sus_scrofa/Info/Index, accessed 20 September 2022). Genes nearest the significant SNPs are list in Tables. We conducted a search of both PubMed and the relevant literature to examine the correlation between the nearest peak SNPs of all the candidate genes and the internal organ weight traits being analyzed.

3. Results and Discussion

3.1. Phenotype Statistics and Heritability Estimation

The descriptive phenotypic statistics and estimated heritabilities (h^2) for analysis of the internal organ weights are listed in Table 1. The weight of internal organs is a crucial indicator of internal organ development and has a significant impact on organ function. In the current study, the average Heart WT, Liver WT, Spleen WT, Lung WT, Kidney WT, and Stomach WT in DLY pigs were 455.57 g, 1763.61 g, 212.54 g, 1020.54 g, 0.41 kg, and 727.68 g, respectively. The estimated heritabilities of Heart WT and Lung WT were the lowest at 0.21 ± 0.04 and 0.28 ± 0.04 , respectively, and all other organ weights had had moderate to high estimated heritabilities, ranging from 0.36 ± 0.04 to 0.49 ± 0.04 . Similar to the results of a previous study, the estimated heritabilities of Heart WT, Liver WT, Spleen WT, and Kidney WT were between 0.35 and 0.54, which were moderate to high estimations [5], indicating that the estimated heritabilities of the weight of an internal organ is generally high in pigs and there is considerable room for improving the genetic contribution through breeding. Furthermore, the coefficients of variation were the lowest for Lung WT and all other traits were relatively high, indicating individual heterogeneity, low trait selection intensity, and high breeding potential.

Table 1. Phenotypic statistics and heritability estimates for Heart WT, Liver WT, Spleen WT, Lung WT, Kidney WT, and Stomach WT.

Trait	N	Mean (\pm SD)	Min	Max	C.V.% ^a	h^2 (\pm SE)
Heart WT	1518	455.57 \pm 78.06	217.1	868.8	17.13	0.21 \pm 0.04
Liver WT	1518	1763.61 \pm 271.83	993.9	2607.7	15.41	0.46 \pm 0.04
Spleen WT	1517	212.54 \pm 47.49	95.9	502.1	22.34	0.49 \pm 0.04
Lung WT	1517	1020.54 \pm 240.10	918.9	2090.4	2.18	0.28 \pm 0.04
Kidney WT	1486	0.41 \pm 0.08	0.17	0.84	19.51	0.36 \pm 0.04
Stomach WT	1518	727.68 \pm 129.97	495.7	1321.7	17.86	0.47 \pm 0.04

^a Coefficient of variation (C.V.).

Moreover, the genetic and phenotypic correlation coefficients among Heart WT, Liver WT, Spleen WT, Lung WT, Kidney WT, and Stomach WT are listed in Table 2. The results revealed moderate to low genetic correlations among the six internal organ weight traits. Heart WT had moderate genetic correlations with Liver WT, Lung WT, and Kidney WT, suggesting that these traits could be improved together in pig breeding programs. On the other hand, Stomach WT showed close to 0 genetic correlations with most of the other traits, indicating Stomach WT traits are less influenced by other traits when they are inherited. Therefore, reasonable breeding strategies need to be designed to improve internal organ

weight traits. The phenotypic correlation results showed that the correlation coefficients between the phenotypes were at moderate to high levels, excluding the low phenotypic correlation coefficients between Lung WT and Liver WT, and Spleen WT and Kidney WT, especially the phenotypic correlation coefficients of Liver WT and Kidney WT were as high as 0.62. When selecting for a certain phenotype in pig breeding, it is advantageous to also consider other related traits.

Table 2. Phenotypic correlations (above the diagonal) and genetic correlations (below the diagonal) among organ weight traits within the DLY population.

	Heart WT	Liver WT	Spleen WT	Lung WT	Kidney WT	Stomach WT
Heart WT	1	0.36	0.33	0.41	0.37	0.49
Liver WT	0.31 ± 0.11	1	0.34	−0.02	0.62	0.38
Spleen WT	0.23 ± 0.11	0.11 ± 0.09	1	0.15	0.33	0.41
Lung WT	0.33 ± 0.13	0.17 ± 0.11	0.04 ± 0.11	1	0.05	0.27
Kidney WT	0.30 ± 0.12	0.33 ± 0.09	0.07 ± 0.10	0.29 ± 0.12	1	0.43
Stomach WT	0.02 ± 0.12	0.03 ± 0.09	0.26 ± 0.09	0.02 ± 0.11	0.03 ± 0.10	1

3.2. Population Structure and LD decay

Population stratification is known to lead to false-positive results in GWASs. To detect potential population stratification, we performed PCA and added the first five principal components to the covariates of the GWAS model to correct for the population structure. Moreover, our previous study showed that the LD decay coefficient of the analyzed DLY pig population with r^2 decayed to 0.2 at a physical distance of 200 kb [24], indicating that the DLY population is diverse with a weak linkage between loci, which facilitates the detection of key SNPs for internal organ weight traits. In addition, Q–Q plots were generated for Heart WT, Liver WT, Spleen WT, Lung WT, Kidney WT, and Stomach WT to further assess population stratification (together with the Manhattan plots: Figure 2). The expansion coefficients (lambda) of the Q–Q plots for all six internal organ weight traits were close to 1, and no overall systematic bias was observed, signifying a negligible effect of the DLY pig group structure on GWASs.

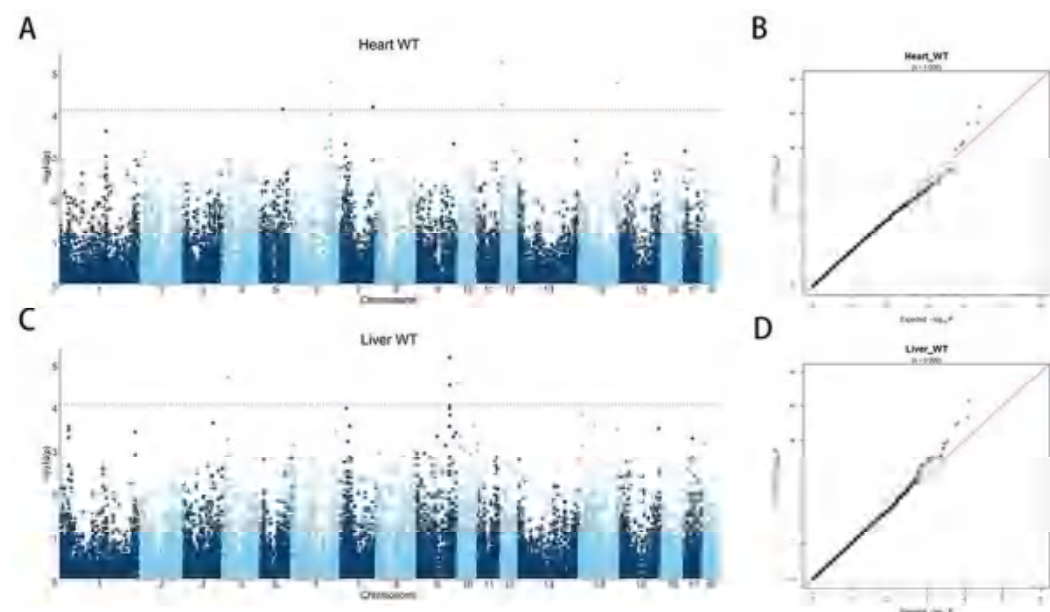


Figure 2. Cont.

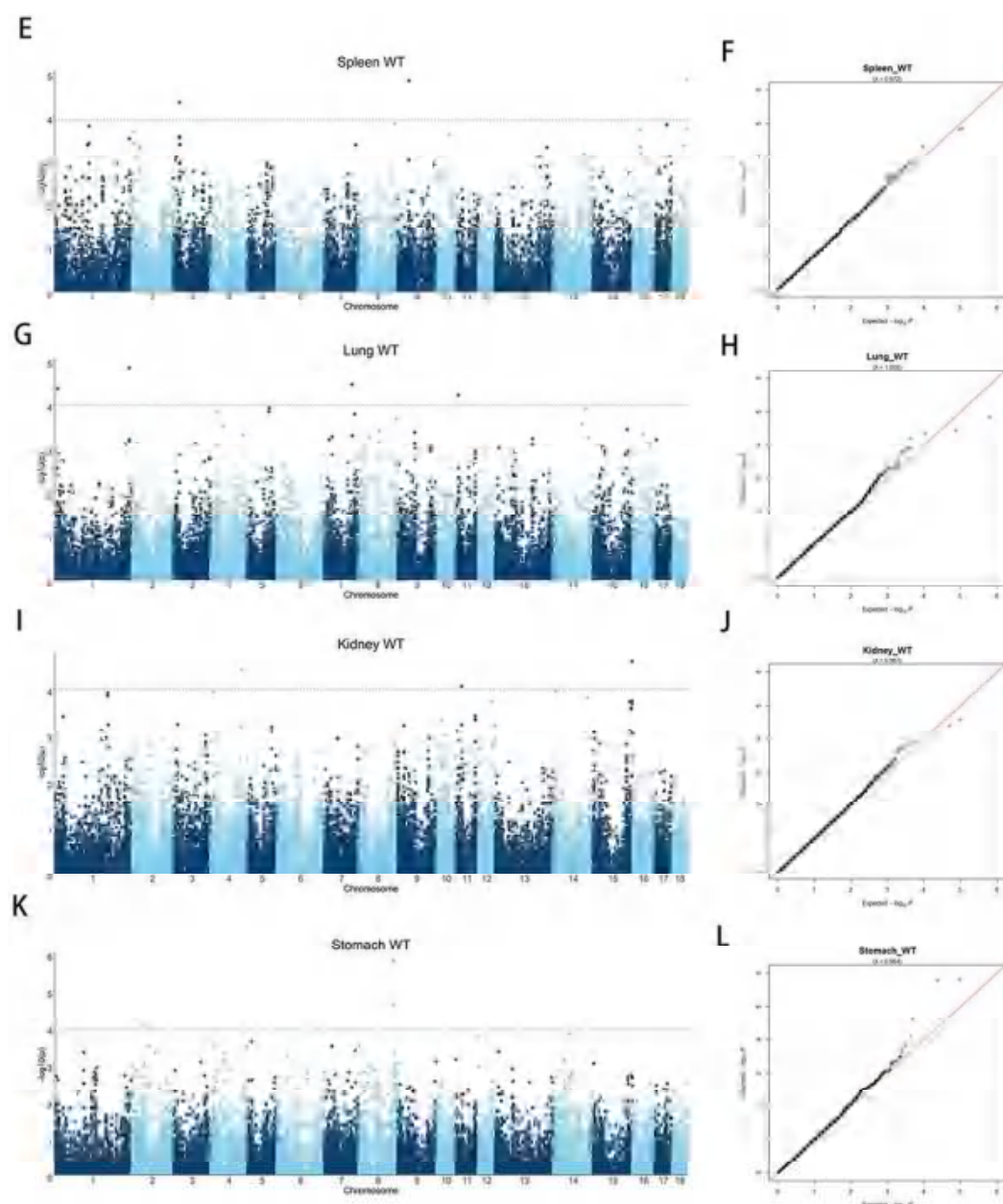


Figure 2. Manhattan and Q-Q plots of internal organ weight traits in the single-trait GWAS. (A) GWAS for Heart WT; (B) Q-Q plot for Heart WT; (C) GWAS for Liver WT; (D) Q-Q plot for Liver WT; (E) GWAS for Spleen WT; (F) Q-Q plot for Spleen WT; (G) GWAS for Lung WT; (H) Q-Q plot for Lung WT; (I) GWAS for Kidney WT; (J) Q-Q plot for Heart WT; (K) GWAS for Stomach WT; (L) Q-Q plot for Stomach WT. The x -axis represents the chromosome, and the y -axis represents the $-\log_{10}(p\text{-value})$ value in the Manhattan plot of the GWAS. The Q-Q plot is plotted with the x -axis representing the actual measured value of $-\log_{10}(p\text{-value})$ and the y -axis representing the observed value of $-\log_{10}(p\text{-value})$ and labeled with the expansion factor lambda (λ).

3.3. Single-Trait GWASs

Single-trait GWASs were performed for the weight of the heart, liver, spleen, lung, kidney, and stomach. The results showed that 6, 4, 3, 4, 3, and 4 SNPs were significantly associated with the weight of each organ, respectively. The results of these single-trait GWASs are presented in Figure 2 and Table 3. Notably, it is the first time that significant SNPs associated with Stomach WT have been identified in pigs. Furthermore, on the basis of the LD decay map, a region of 200 kb before and after the key SNPs was defined as a region to screen for candidate genes [24]. For heart weight, six significant SNPs were identified,

located on SSC5, 6, 7, 12, and 14. These six SNPs surpassed the significance threshold of 1.01×10^{-4} . Figure 2B shows an expansion coefficient lambda (λ) of 1.006. Details of the significant SNPs are listed in Table 3. The most significant SNP, WU_10.2_12_6703865 on SSC12, explains 1.70% of the phenotypic variation and is about 44 kb downstream of the *CD300LB* gene. The *CD300LB* gene is a triggering receptor expressed on bone marrow cells that regulates the cytosolic process of bone marrow cells [27], and the *CD300LB* protein stimulated by T cells regulates *DNMT3A* mutation and alters immune cells in heart failure [28].

Table 3. Significant SNPs and candidate genes for Heart WT, Liver WT, Spleen WT, Lung WT, Kidney WT, and Stomach WT in single-trait GWASs.

Trait	SSC	SNP	Position (bp)	MAF	<i>p</i> -Value	PEV (%) ^a	Candidate Gene	Distance
Heart WT	6	WU_10.2_6_126961053	137,008,535	0.257	1.88×10^{-5}	2.09%	<i>ST6GALNAC3</i>	Within
	14	6_43731895	132,228,312	0.407	1.99×10^{-5}	1.71%	<i>HTRA1</i>	124,717
	12	WU_10.2_12_6703865	6,682,110	0.389	6.40×10^{-5}	1.70%	<i>CD300LB</i>	44,143
	5	ALGA0032998	76,317,972	0.264	8.30×10^{-5}	1.36%	<i>ANO6</i>	Within
	7	WU_10.2_7_116585612	110,088,932	0.242	7.42×10^{-5}	1.10%	<i>KCNK10</i>	−29,135
	12	12_5381300	5,426,010	0.319	6.50×10^{-5}	0.53%	<i>CDK3</i>	Within
Liver WT	4	WU_10.2_4_20570494	19,550,555	0.249	2.15×10^{-5}	2.11%	<i>CCN3</i>	−35,758
	9	H3GA0028070	113,152,352	0.09	7.54×10^{-5}	2.10%	<i>TPK1</i>	Within
	9	ASGA0044340	113,140,343	0.118	3.51×10^{-5}	0.82%	<i>TPK1</i>	Within
	10	WU_10.2_10_3469625	1,753,591	0.476	3.08×10^{-5}	0.43%	<i>RGS21</i>	Within
Spleen WT	18	ALGA0098928	54,993,603	0.328	1.39×10^{-5}	2.22%	<i>POU6F2</i>	Within
	3	ALGA0105765	20,525,651	0.216	4.74×10^{-5}	0.78%	<i>HS3ST4</i>	−19,582
	9	WU_10.2_9_45824613	40,989,995	0.399	1.47×10^{-5}	0.58%	<i>TTC12</i>	Within
Lung WT	11	ALGA0060656	8,759,687	0.231	6.24×10^{-5}	2.76%	<i>FRY</i>	Within
	1	ALGA0110225	266,708,292	0.278	1.46×10^{-5}	1.81%	<i>PBX3</i>	97,646
	7	ASGA0035515	98,022,168	0.04	3.57×10^{-5}	1.43%	<i>YLP1M1</i>	Within
	1	MARC0089438	11,012,474	0.271	4.46×10^{-5}	0.69%	/	/
Kidney WT	15	WU_10.2_15_153747936	138,955,316	0.369	2.61×10^{-5}	1.97%	/	/
	4	WU_10.2_4_119054114	108,872,194	0.478	3.93×10^{-5}	1.62%	<i>RAP1A</i>	122,469
	11	WU_10.2_11_20231427	19,917,312	0.221	8.86×10^{-5}	0.88%	<i>SUCLA2</i>	Within
Stomach WT	8	ALGA0106192	124,443,641	0.064	1.54×10^{-5}	2.62%	<i>UNC5C</i>	Within
	8	MARC0052872	124,421,940	0.063	1.61×10^{-5}	2.53%	<i>UNC5C</i>	Within
	8	ASGA0101191	124,548,240	0.081	2.40×10^{-5}	2.14%	<i>BMP1R1B</i>	Within
	2	MARC0018316	46,014,239	0.46	8.35×10^{-5}	0.74%	<i>ARNTL</i>	Within

^a The percentage of phenotypic variance explained by each SNP.

For liver weight, four significant SNPs were detected on SSC4, SSC9, and SSC10 with a λ of 0.999 (Figure 2C,D and Table 3). These four SNPs surpassed the significance threshold of 1.04×10^{-4} . The top SNP, H3GA0028070, accounted for 2.10% of the phenotypic variance and is located within the *TPK1* gene. A significant SNP, named ASGA0044340, 12 kb upstream of H3GA0028070, also located on *TPK1*, explained 0.82% of the phenotypic variation. According to reports, *TPK1* is a cofactor of certain enzymes associated with the glycolysis and energy production pathways. It is involved in the metabolism of water-soluble vitamins and cofactors and the thiamine metabolic pathway, and mutations in *TPK1* can cause thiamine metabolic dysfunction syndrome [29]. In addition, knockdown of this gene can lead to glycogen storage dysfunction [30]. However, no studies have shown *TPK1* to be directly associated with liver development and weight in pigs.

The GWAS results of Spleen WT identified three significant SNPs, located on SSC3, SSC9, and SSC18, with a λ of 0.972 (Figure 2E,F and Table 3). All three SNPs surpassed the threshold of significance ($p < 1.20 \times 10^{-4}$). The top SNP, ALGA0098928, explained 2.22% of the phenotypic variation and is located within *POU6F2*. *POU6F2* is a suppressor associated with nephroblastoma (WT) that regulates cell proliferation and specific differentiation [31]. According to the RT-qPCR results, the expression of *POU6F2* is associated with renal morphogenesis [32], suggesting that *POU6F2* may be closely associated with spleen weight traits.

For lung weight, four significant SNPs were detected on SSC1, SSC7, and SSC11 with a λ of 1.008 (Figure 2G,H and Table 3). These four SNPs surpassed the significance threshold of 1.04×10^{-4} . An SNP named ALGA0110225 explained 1.81% of the phenotypic variance and is located 97 kb downstream of *PBX3*, indicating that *PBX3* and ALGA0110225 may both play a role in Lung WT. A literature review revealed that *PBX3* is directly regulated by targeting *NBPF10*, *miR-144*, and *miR-224*, which are directly associated with lung cancer cell proliferation [33]. In addition, overexpression of *PBX3* promotes the proliferation of A549 cells (lung cancer histiocytes) [34]. Therefore, we believe *PBX3* to be a promising candidate gene for influencing Lung WT, and the regulatory mechanism needs further investigation.

We performed GWASs with the Kidney WT trait in DLY pigs and detected three SNPs that were above the significance threshold ($p < 1.03 \times 10^{-4}$) (Figure 2I and Table 3). Figure 2J shows that the lambda is 0.987. WU_10.2_15_153747936 on SSC15 explains 1.97% of the phenotypic variation and is located 341 kb downstream of the *HDAC4* gene. Because the LD decay distance is 200 kb, it follows that *HDAC4* may not have a significant effect on Kidney WT traits.

QTL and significant SNPs have not been previously reported in relation to Stomach WT. Thus, this study is the first GWAS on pig Stomach WT. Herein, four significant SNPs were identified for the DLY pigs that were above the significance threshold of $p < 1.17 \times 10^{-4}$ (Figure 2K and Table 3). Of the four significant SNPs, three SNPs were simultaneously located on SSC8 and both MARC0052872 and ALGA0106192 were located within the *UNC5C* gene. Furthermore, the distance between MARC0052872 and ALGA0106192 was only 21 kb, explaining 2.53% and 2.62% of the phenotypic variation, respectively. ASGA0101191 was 100 kb away from the aforementioned SNPs with a phenotypic variation value of 2.14% and was located within *BMPR1B*. A literature review revealed that *UNC5C* plays a dominant role in netrin-1/*UNC5C*-mediated axonal rejection [35] and that its promoter region sequence binds to *p53* and acts as a target of *p53* to regulate apoptosis [36]. As regards the *BMPR1B* gene, it has been shown that the BMP family is expressed in the early organ and tissue formation during mouse embryonic development [37]. However, neither the *BMPR1B* nor the *UNC5C* gene is directly associated with internal organ weight traits.

The above GWAS results show that none of the SNPs associated with internal organ weight overlapped with those previously reported QTL documented in the pig QTL database [1]. This may have been due to the fact that most studies focused on the breeding of native Chinese pigs, and fewer studies were conducted on DLY three-way crossbred commercial populations with significant breed differences. Moreover, the significant SNPs did not overlap in the six traits, i.e., none of the SNPs were polymorphic, which may be related to the low genetic and phenotypic correlation between traits and the low density of genetic markers, which was further verified by multi-trait GWASs.

3.4. Haplotype Block Analysis

Figure 3 shows the LD pattern of significant SNPs associated with Stomach WT. In this study, multiple SNPs associated with Stomach WT were in close proximity to each other, with two significant SNPs on SSC8, which is located in a 21 kb region within the *UNC5C* and *BMPR1B* genes (the gene function is described above). The insufficient density of 50K microarray markers resulted in a low number of SNPs with linkage disequilibrium, which limited the resolution of the genetic architecture of key SNPs for the trait to some extent.

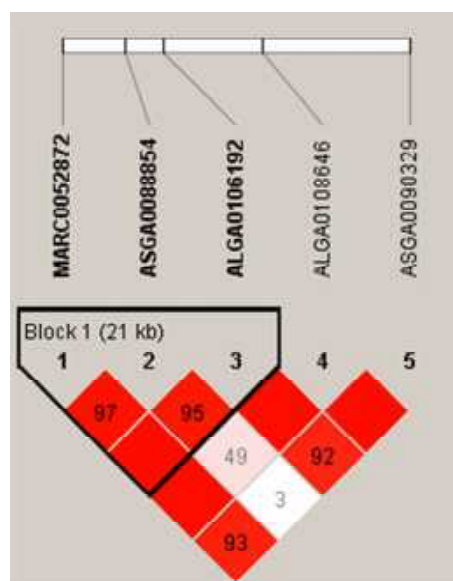


Figure 3. Linkage disequilibrium blocks of important SNPs on SSC8 in DLY pigs. The value in the box is the degree of linkage disequilibrium between SNPs (r^2).

3.5. Multi-Trait GWASs

In order to improve the statistical effect, multi-trait GWASs were individually performed for each SNP by combining the joint analysis of six internal organ weight traits. This revealed the genetic factors with significant interactions among different traits in the same individual under the same environment. Manhattan plots of the multi-trait GWASs are shown in Figure 4.

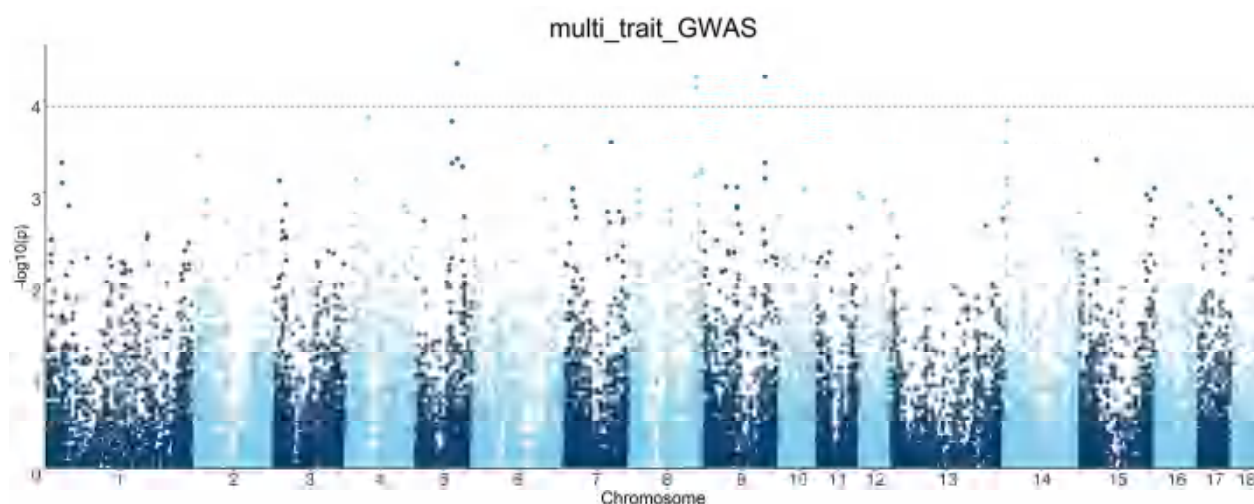


Figure 4. Manhattan plots of multi-trait GWAS results for six internal organ weight traits in the DLY population. The x-axis represents the chromosome, and the y-axis represents the $-\log_{10}(p\text{-value})$ value in the Manhattan plot of the GWAS. The dashed line indicates the FDR-corrected threshold.

The multi-trait GWASs combining six internal organ weight traits identified four significant SNPs with polymorphisms affecting the phenotypes, ALGA0032998, H3GA0028070, MARC0052872, and ALGA0106192 (Figure 5 and Table 4). SNP ALGA0032998 explained 1.36% of the phenotypic variation and is located within the *ANO6* gene. The overexpression of *CCR7* was observed to enhance the migration of BxPC-3 cells under the induction of the *ANO6* gene, which is a potential mediator of *ANO6* expression through the *ERK* signaling pathway. This promotion of migration was also seen in pancreatic ductal adenocarcinoma cells [38]. The single-trait GWAS described the effects of three SNPs (H3GA0028070,

MARC0052872, and ALGA0106192) located on *TPK1* and *UNC5C* genes on the weight of the liver and stomach. These SNPs were not only significant in the single-trait GWAS but were also found to be simultaneously associated with the weight of all six internal organs, suggesting that these four SNPs have pleiotropic effects. Furthermore, no additional SNPs, independent of the single-trait GWAS results, were found. Similar results were previously reported by Guo et al. [39], in which no additional SNPs, independent of the single-trait GWAS results, were detected in the multi-trait GWASs for backfat thickness, carcass weight, and body weight in the DLY and Duroc populations. The reasons for this situation are manifold. For example, the complexity of the genetic architecture of the internal organ weight trait and the low marker density result in a low number of SNPs reaching significant levels. This renders LD detection insufficient and increases the difficulty of screening for co-dominant SNP or QTL regions. Thus, a larger sample population and a higher marker density are required to screen for loci associated with internal organ weight.

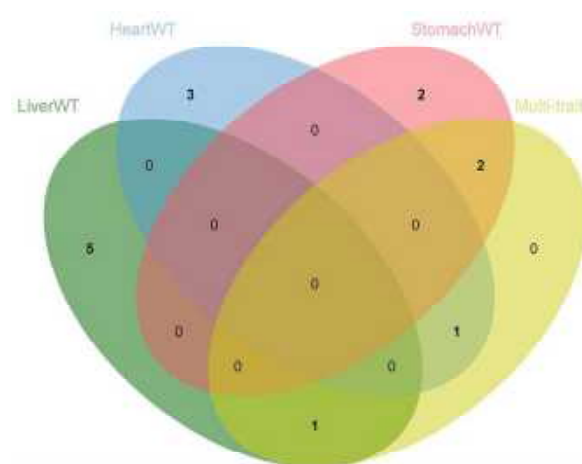


Figure 5. Venn diagram showing the distribution of SNPs in the single-trait GWASs and multi-trait GWASs, highlighting the role of pleiotropic SNPs in multiple traits.

Table 4. Significant SNPs and candidate genes for Heart WT, Liver WT, Spleen WT, Lung WT, Kidney WT, and Stomach WT in multi-trait GWASs.

SSC	SNP	Position (bp)	MAF	<i>p</i> -Value	PEV (%) ^a	Candidate Gene	Distance
8	ALGA0106192	124,443,641	0.064	7.29×10^{-5}	2.62%	<i>UNC5C</i>	Within
8	MARC0052872	124,421,940	0.063	5.61×10^{-5}	2.53%	<i>UNC5C</i>	Within
9	H3GA0028070	113,152,352	0.088	5.51×10^{-5}	2.10%	<i>TPK1</i>	Within
5	ALGA0032998	76,317,972	0.266	3.88×10^{-5}	1.36%	<i>ANO6</i>	Within

^a The percentage of phenotypic variance explained by each SNP.

4. Conclusions

In this study, we conducted single-trait and multi-trait GWASs on the internal organ weights of 1518 DLY pigs. A total of 24 significant SNPs were detected in the single-trait GWAS results for six internal organ weight traits. The four significant pleiotropic SNPs identified via multi-trait GWASs were associated with six internal organ weight traits, confirming the results of the single-trait GWASs and improving our ability to reveal the genetic architecture of organ weight traits. *TPK1*, *POU6F2*, *PBX3*, *UNC5C*, and *BMPRI1B* were highlighted as potential genes responsible for differences in Liver WT, Spleen WT, Lung WT, and Stomach WT among individuals according to their gene functions. In summary, the results of this study contribute to our understanding of the genetics of internal organ weight traits in DLY pigs by assigning higher weights to relevant SNPs and key genes in the genome.

Author Contributions: Conceptualization, E.Z. and G.C.; resources, E.Z., G.C. and Z.W.; investigation, Y.Y., S.Z., D.R., Y.Q., S.W. and J.W.; software and validation, X.L., Z.Z. and J.W.; formal analysis, X.L., J.W. and Z.Z.; data curation, Z.Z., S.W., J.W. and Y.Y.; writing—original draft preparation, X.L. and J.W.; writing—review and editing, E.Z., G.C., Z.W. and J.Y.; visualization, J.Y. and Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Key Technologies R&D Program of Guangdong Province project (2022B0202090002) and the Project of Swine Innovation Team in Guangdong Modern Agricultural Research System (2022KJ126).

Institutional Review Board Statement: All animals used in this study were treated in accordance with the guidelines for the use of laboratory animals of the Ministry of Agriculture of China and with the approval of South China Agricultural University (Guangzhou, China), No. 2018F089.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. An, B.; Xia, J.; Chang, T.; Wang, X.; Miao, J.; Xu, L.; Zhang, L.; Gao, X.; Chen, Y.; Li, J.; et al. Genome-wide association study identifies loci and candidate genes for internal organ weights in Simmental beef cattle. *Physiol. Genom.* **2018**, *50*, 523–531. [[CrossRef](#)] [[PubMed](#)]
2. Mader, C.J.; Montanholi, Y.R.; Wang, Y.J.; Miller, S.P.; Mandell, I.B.; McBride, B.W.; Swanson, K.C. Relationships among measures of growth performance and efficiency with carcass traits, visceral organ mass, and pancreatic digestive enzymes in feedlot cattle. *J. Anim. Sci.* **2009**, *87*, 1548–1557. [[CrossRef](#)]
3. Mubbunu, L.; Bowa, K.; Petrenko, V.; Silitongo, M. Correlation of Internal Organ Weights with Body Weight and Body Height in Normal Adult Zambians: A Case Study of Ndola Teaching Hospital. *Anat. Res. Int.* **2018**, *2018*, 4687538. [[CrossRef](#)]
4. Hu, Z.L.; Park, C.A.; Reecy, J.M. Bringing the Animal QTLdb and CorrDB into the future: Meeting new challenges and providing updated services. *Nucleic Acids Res.* **2022**, *50*, D956–D961. [[CrossRef](#)] [[PubMed](#)]
5. He, Y.; Li, X.; Zhang, F.; Su, Y.; Hou, L.; Chen, H.; Zhang, Z.; Huang, L. Multi-breed genome-wide association study reveals novel loci associated with the weight of internal organs. *Genet. Sel. Evol.* **2015**, *47*, 87. [[CrossRef](#)] [[PubMed](#)]
6. Zhang, J.; Xiong, Y.; Zuo, B.; Lei, M.; Jiang, S.; Li, F.; Zheng, R.; Li, J.; Xu, D. Detection of quantitative trait loci associated with several internal organ traits and teat number trait in a pig population. *J. Genet. Genom.* **2007**, *34*, 307–314. [[CrossRef](#)] [[PubMed](#)]
7. Liu, X.; Wang, L.; Liang, J. Genome-Wide Association Study for Certain Carcass Traits and Organ Weights in a Large White×Minzhu Intercross Porcine Population. *J. Integr. Agric.* **2014**, *13*, 2721–2730. [[CrossRef](#)]
8. Wei, W.H.; Duan, Y.; Haley, C.S.; Ren, J.; de Koning, D.J.; Huang, L.S. High throughput analyses of epistasis for swine body dimensions and organ weights. *Anim. Genet.* **2011**, *42*, 15–21. [[CrossRef](#)]
9. Casale, F.P.; Rakitsch, B.; Lippert, C.; Stegle, O. Efficient set tests for the genetic analysis of correlated traits. *Nat. Methods* **2015**, *12*, 755–758. [[CrossRef](#)]
10. Porter, H.F.; O'Reilly, P.F. Multivariate simulation framework reveals performance of multi-trait GWAS methods. *Sci. Rep.* **2017**, *7*, 38837. [[CrossRef](#)]
11. Broadaway, K.A.; Cutler, D.J.; Duncan, R.; Moore, J.L.; Ware, E.B.; Jhun, M.A.; Bielak, L.F.; Zhao, W.; Smith, J.A.; Peyser, P.A.; et al. A Statistical Approach for Testing Cross-Phenotype Effects of Rare Variants. *Am. J. Hum. Genet.* **2016**, *98*, 525–540. [[CrossRef](#)] [[PubMed](#)]
12. Hackinger, S.; Zeggini, E. Statistical methods to detect pleiotropy in human complex traits. *Open Biol.* **2017**, *7*, 170125. [[CrossRef](#)] [[PubMed](#)]
13. Zhou, S.; Ding, R.; Zhuang, Z.; Zeng, H.; Wen, S.; Ruan, D.; Wu, J.; Qiu, Y.; Zheng, E.; Cai, G.; et al. Genome-Wide Association Analysis Reveals Genetic Loci and Candidate Genes for Chest, Abdominal, and Waist Circumferences in Two Duroc Pig Populations. *Front. Vet. Sci.* **2021**, *8*, 807003. [[CrossRef](#)] [[PubMed](#)]
14. An, B.; Xu, L.; Xia, J.; Wang, X.; Miao, J.; Chang, T.; Song, M.; Ni, J.; Xu, L.; Zhang, L.; et al. Multiple association analysis of loci and candidate genes that regulate body size at three growth stages in Simmental beef cattle. *BMC Genet.* **2020**, *21*, 32. [[CrossRef](#)]
15. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
16. Yang, J.; Lee, S.H.; Goddard, M.E.; Visscher, P.M. GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **2011**, *88*, 76–82. [[CrossRef](#)]
17. Zhou, X.; Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **2012**, *44*, 821–824. [[CrossRef](#)]

18. Zhou, X.; Stephens, M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* **2014**, *11*, 407–409. [\[CrossRef\]](#)
19. Yan, G.; Guo, T.; Xiao, S.; Zhang, F.; Xin, W.; Huang, T.; Xu, W.; Li, Y.; Zhang, Z.; Huang, L. Imputation-Based Whole-Genome Sequence Association Study Reveals Constant and Novel Loci for Hematological Traits in a Large-Scale Swine F(2) Resource Population. *Front. Genet.* **2018**, *9*, 401. [\[CrossRef\]](#)
20. Glickman, M.E.; Rao, S.R.; Schultz, M.R. False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *J. Clin. Epidemiol.* **2014**, *67*, 850–857. [\[CrossRef\]](#)
21. Barrett, J.C.; Fry, B.; Maller, J.; Daly, M.J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **2005**, *21*, 263–265. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Ding, R.; Quan, J.; Yang, M.; Wang, X.; Zheng, E.; Yang, H.; Fu, D.; Yang, Y.; Yang, L.; Li, Z.; et al. Genome-wide association analysis reveals genetic loci and candidate genes for feeding behavior and eating efficiency in Duroc boars. *PLoS ONE* **2017**, *12*, e0183244. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Zaitlen, N.; Kraft, P. Heritability in the genome-wide association era. *Hum. Genet.* **2012**, *131*, 1655–1664. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Zhuang, Z.; Wu, J.; Xu, C.; Ruan, D.; Qiu, Y.; Zhou, S.; Ding, R.; Quan, J.; Yang, M.; Zheng, E.; et al. The Genetic Architecture of Meat Quality Traits in a Crossbred Commercial Pig Population. *Foods* **2022**, *11*, 3143. [\[CrossRef\]](#)
25. Zhuang, Z.; Ding, R.; Peng, L.; Wu, J.; Ye, Y.; Zhou, S.; Wang, X.; Quan, J.; Zheng, E.; Cai, G.; et al. Genome-wide association analyses identify known and novel loci for teat number in Duroc pigs using single-locus and multi-locus models. *BMC Genom.* **2020**, *21*, 344. [\[CrossRef\]](#)
26. Durinck, S.; Spellman, P.T.; Birney, E.; Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **2009**, *4*, 1184–1191. [\[CrossRef\]](#)
27. Voss, O.H.; Tian, L.; Murakami, Y.; Coligan, J.E.; Krzewski, K. Emerging role of CD300 receptors in regulating myeloid cell efferocytosis. *Mol. Cell. Oncol.* **2015**, *2*, e964625. [\[CrossRef\]](#)
28. Abplanalp, W.T.; Cremer, S.; John, D.; Hoffmann, J.; Schuhmacher, B.; Merten, M.; Rieger, M.A.; Vasa-Nicotera, M.; Zeiher, A.M.; Dimmeler, S. Clonal Hematopoiesis-Driver DNMT3A Mutations Alter Immune Cells in Heart Failure. *Circ. Res.* **2021**, *128*, 216–228. [\[CrossRef\]](#)
29. Li, X.; Huang, Z.; Chen, Y.; Sun, X.; Yi, Z.; Xie, J.; Yu, X.; Chen, H.; Zhong, J. Case report of two affected siblings in a family with thiamine metabolism dysfunction syndrome 5: A rare, but treatable neurodegenerative disease. *BMC Neurol.* **2022**, *22*, 373. [\[CrossRef\]](#)
30. Giacometti, R.; Kronberg, F.; Biondi, R.M.; Passeron, S. Catalytic isoforms Tpk1 and Tpk2 of *Candida albicans* PKA have non-redundant roles in stress response and glycogen storage. *Yeast* **2009**, *26*, 273–285. [\[CrossRef\]](#)
31. Di Renzo, F.; Doneda, L.; Menegola, E.; Sardella, M.; De Vecchi, G.; Collini, P.; Spreafico, F.; Fossati-Bellani, F.; Giavini, E.; Radice, P.; et al. The murine Pou6f2 gene is temporally and spatially regulated during kidney embryogenesis and its human homolog is overexpressed in a subset of Wilms tumors. *J. Pediatr. Hematol. Oncol.* **2006**, *28*, 791–797. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Xu, G.; Zhu, H.; Xu, J.; Wang, Y.; Zhang, Y.; Zhang, M.; Zhu, D. Long non-coding RNA POU6F2-AS2 promotes cell proliferation and drug resistance in colon cancer by regulating miR-377/BRD4. *J. Cell. Mol. Med.* **2020**, *24*, 4136–4149. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Zhao, X.; Song, X.; Zhao, Y. circNBP10/miR-224 Axis Regulates PBX3 to Promote the Malignant Progression of Lung Cancer. *J. Oncol.* **2022**, *2022*, 2832920. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Li, D.; Li, H.; Yang, Y.; Kang, L. Long Noncoding RNA Urothelial Carcinoma-Associated 1 Promotes the Proliferation and Metastasis of Human Lung Tumor Cells by Regulating MicroRNA-144. *Oncol. Res.* **2018**, *26*, 537–546. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Shao, Q.; Yang, T.; Huang, H.; Alarmanazi, F.; Liu, G. Uncoupling of UNC5C with Polymerized TUBB3 in Microtubules Mediates Netrin-1 Repulsion. *J. Neurosci.* **2017**, *37*, 5620–5633. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Miyamoto, Y.; Futamura, M.; Kitamura, N.; Nakamura, Y.; Baba, H.; Arakawa, H. Identification of UNC5A as a novel transcriptional target of tumor suppressor p53 and a regulator of apoptosis. *Int. J. Oncol.* **2010**, *36*, 1253–1260. [\[CrossRef\]](#)
37. Danesh, S.M.; Villasenor, A.; Chong, D.; Soukup, C.; Cleaver, O. BMP and BMP receptor expression during murine organogenesis. *Gene Expr. Patterns* **2009**, *9*, 255–265. [\[CrossRef\]](#)
38. Wang, L.; Zhao, X.Y.; Zhu, J.S.; Chen, N.W.; Fan, H.N.; Yang, W.; Guo, J.H. CCR7 regulates ANO6 to promote migration of pancreatic ductal adenocarcinoma cells via the ERK signaling pathway. *Oncol. Lett.* **2018**, *16*, 2599–2605. [\[CrossRef\]](#)
39. Guo, Y.; Qiu, H.; Xiao, S.; Wu, Z.; Yang, M.; Yang, J.; Ren, J.; Huang, L. A genome-wide association study identifies genomic loci associated with backfat thickness, carcass weight, and body weight in two commercial pig populations. *J. Appl. Genet.* **2017**, *58*, 499–508. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Treatment of Donor Cells with Oxidative Phosphorylation Inhibitor CPI Enhances Porcine Cloned Embryo Development

Jinping Cao ^{1,2,3,4,5,6,†}, Yazheng Dong ^{1,2,3,4,5,6,†}, Zheng Li ^{1,2,3,4,5,6}, Shunbo Wang ^{1,2,3,4,5,6}, Zhenfang Wu ^{1,2,3,4,5,6}, Enqin Zheng ^{1,2,3,4,5,6,*} and Zicong Li ^{1,2,3,4,5,6,*}

- ¹ National Engineering Research Center for Breeding Swine Industry, South China Agricultural University, Guangzhou 510642, China; caojinping2021@scau.edu.cn (J.C.); dongyazheng@stu.scau.edu.cn (Y.D.); lizheng2022@stu.scau.edu.cn (Z.L.); 03028301@stu.scau.edu.cn (S.W.); wzfemail@163.com (Z.W.)
 - ² State Key Laboratory of Swine and Poultry Breeding Industry, South China Agricultural University, Guangzhou 510642, China
 - ³ National and Local Joint Engineering Research Center for Livestock and Poultry Breeding Industry, South China Agricultural University, Guangzhou 510642, China
 - ⁴ Department of Animal Genetics, Breeding and Reproduction, College of Animal Science, South China Agricultural University, Guangzhou 510642, China
 - ⁵ Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, South China Agricultural University, Guangzhou 510642, China
 - ⁶ Gene Bank of Guangdong Local Livestock and Poultry, South China Agricultural University, Guangzhou 510642, China
- * Correspondence: eqzheng@scau.edu.cn (E.Z.); lizicong@scau.edu.cn (Z.L.)
† These authors contributed equally to this work.

Simple Summary: The low developmental efficiency of cloned pig embryos limits the application of the pig cloning technique. In this study, a small molecule drug called CPI was added to the culture medium of nuclear donor cells to enhance the developmental capacity of subsequent cloned pig embryos. CPI treatment of nuclear donor cells changed the cellular energy metabolism status from oxidative phosphorylation to glycolysis, and significantly improved the developmental competence of subsequent cloned pig embryos. This study provides a simple approach to facilitate the development and application of pig cloning technology.



Citation: Cao, J.; Dong, Y.; Li, Z.; Wang, S.; Wu, Z.; Zheng, E.; Li, Z. Treatment of Donor Cells with Oxidative Phosphorylation Inhibitor CPI Enhances Porcine Cloned Embryo Development. *Animals* **2024**, *14*, 1362. <https://doi.org/10.3390/ani14091362>

Received: 29 March 2024
Revised: 27 April 2024
Accepted: 29 April 2024
Published: 30 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Somatic cell nuclear transfer (SCNT) technology holds great promise for livestock industry, life science and human biomedicine. However, the development and application of this technology is limited by the low developmental potential of SCNT embryos. The developmental competence of cloned embryos is influenced by the energy metabolic status of donor cells. The purpose of this study was to investigate the effects of CPI, an oxidative phosphorylation inhibitor, on the energy metabolism pathways of pig fibroblasts and the development of subsequent SCNT embryos. The results showed that treatment of porcine fibroblasts with CPI changed the cellular energy metabolic pathways from oxidative phosphorylation to glycolysis and enhanced the developmental ability of subsequent SCNT embryos. The present study establishes a simple, new way to improve pig cloning efficiency, helping to promote the development and application of pig SCNT technology.

Keywords: SCNT; CPI; oxidative phosphorylation; glycolysis

1. Introduction

Somatic cell nuclear transfer (SCNT), also called cloning, enables the development of reconstructed embryos into complete individuals possessing identical nuclear genetic material to that of their donor cells. This technology holds significant application value in animal husbandry, life sciences, and human biomedicine. SCNT can be used to propagate excellent breeding stock, including top bulls [1], Duroc boars [2], Pietran boars [3], and help to save endangered animals such as European argali [4], African wildcats [5], and

eastern Guangdong black pigs. SCNT is also employed to generate various genetically modified animals for use as human disease models [6,7], human xenotransplantation organ donors [8,9], and drug synthesis bioreactors [10,11]. Moreover, SCNT technology facilitates the generation of nuclear transfer embryonic stem cells (ntESC) for applications in human therapeutic cloning [12,13].

Nevertheless, the current efficiency of SCNT technology remains very low. Specifically, the development of pig SCNT embryos into blastocysts under in vitro culture conditions achieves an efficiency of approximately 20–30% [14–16], while in cattle, this efficiency ranges from 20% to 50% [17–19]. Furthermore, the full-term developmental rate of cloned pig embryos after transplantation into the reproductive tract of surrogate mothers is approximately 0.5–3% (calculated as the number of cloned animals born divided by the number of 1–2 cell stage cloned embryos transplanted) [20]. Cattle SCNT embryos exhibit a slightly higher efficiency, with a birth rate of about 10% (calculated as the number of cloned animals born divided by the number of blastocyst stage embryos transplanted) [21]. These considerably low developmental rates pose limitations to the development and application of SCNT technology.

The energy utilization status of donor cells affects the subsequent development of SCNT embryos. Different cell types rely on different metabolic pathways to generate energy. Embryonic stem cells (ESCs) induced pluripotent stem cells (iPSCs), pluripotent stem cells (PSCs), and highly proliferative cells (such as endothelial cells, epithelial cells, and immune cells) predominantly depend on aerobic glycolysis for energy supply [22,23]. Notably, cells in early embryos, being stem cells, also primarily employ glycolysis to generate ATP for cell survival and proliferation [24]. In contrast, other types of somatic cells typically utilize mitochondrial oxidative phosphorylation as the primary source of energy production [25]. Evidence indicates that the transition of cellular energy metabolism pathways from oxidative phosphorylation to glycolysis is essential for reverting cells to an undifferentiated state and maintaining stemness [26]. Comparative analyses of the transcriptome of human iPSCs and fibroblasts reveal upregulation of glycolysis-related enzymes, and a shift of energy metabolism towards glycolysis in iPSCs [27]. The process of reprogramming mouse fibroblasts into totipotent stem cells also involves a switch from oxidative phosphorylation to glycolysis, resulting in reduced cellular oxygen consumption and increased lactate production [28].

The alterations in energy metabolism pathways are typically concomitant with the reprogramming of cell fates [29]. The changes in donor cell metabolism pathways have the potential to enhance the subsequent developmental efficiency of cloned pig embryos. A study has demonstrated that buffalo fetal fibroblasts with elevated expression levels of glycolytic enzyme genes exhibit a higher cloning efficiency [30]. Furthermore, treating buffalo fetal fibroblasts with glycolysis inducer PS48 resulted in the increased production of intracellular lactic acid and a consequential enhancement in the development of cloned embryos [30]. The treatment of porcine fibroblasts with 100 μ M CoCl_2 (a glycolysis inducer) led to the increased expression of glycolytic enzyme genes and the enhanced development of cloned embryos [31].

To investigate the influence of modifying donor cell energy metabolism pathways on the subsequent development of pig SCNT embryos, this study employed the oxidative phosphorylation inhibitor CPI to treat porcine fetal fibroblasts (PFFs), donor cells for cloned embryos. The energy metabolic status and global gene expression pattern were compared between CPI-treated and solvent-incubated negative control (NC) PFFs. The developmental competence of SCNT embryos generated from CPI-treated and NC PFFs was also compared.

2. Materials and Methods

2.1. Porcine Fetal Fibroblast Culture and Passaging

PFFs were obtained from Livestock Germplasm Resource Center, South China Agricultural University (Catalog No. B13XEHTF020200109, small-eared flower breed). The cells were cultured in a 37 °C, 5% CO₂ thermostatic incubator. Upon achieving a cell growth confluence of 80~90%, the cells were ready for passaging. A 1 mL solution containing 0.25% Trypsin-EDTA was added, and the cells were incubated for 2 min. The digestion process was terminated by adding 2 mL of complete medium. Subsequently, centrifugation at 800 rpm for 5 min was performed, the supernatant was discarded, and the cells were transferred to two 100 mm culture dishes for ongoing cultivation.

2.2. CPI Treatment of Donor Cells

The concentration of CPI used for incubating donor cells was chosen according to a previously reported study [32]. Five mg of CPI (MedChem Express Biotechnology, Inc., Monmouth Junction, NJ, USA, item number 95809-78-2) was dissolved in 1.2867 mL of DMSO, preparing a mother liquor with a final concentration of 10 mM. Subsequently, aliquots of 10 µL, 50 µL, and 100 µL of the CPI mother liquor were added to 10 mL of the DMEM complete medium; the final concentrations of the resulting working solution were 10 µM, 50 µM and 100 µM, respectively.

2.3. Reverse Transcription and qPCR

PFF RNA was extracted utilizing the conventional Trizol method, and the reverse transcription was prepared using Evo M-mLV Reverse Transcription Premix Kit II (Guangzhou Ruizhen Biotechnology Co., Ltd., Guangzhou, China). ChamQ Universal SYBR qPCR Master Mix kit (Nanjing Novelty Biotechnology Co., Nanjing, China) was used for qPCR. Primers and their sequences are shown in Table 1. PCR reactions were performed using a Quant Studio 7 Flex system (Thermo Fisher Scientific, Waltham, MA, USA) according to the parameters in Table 2. The qPCR was performed following a previously reported study [33]. The relative expression of genes was calculated using the $2^{-\Delta\Delta C_t}$ method. The data of each group were from three independent samples and each sample was measured with three technical replicates. A melting curve analysis was added at the end of the amplification procedure to confirm the specificity of the amplification.

Table 1. Primer sequence information.

Gene	Primer Sequences (5'–3')	GenBank ID
β-actin (reference gene)	F: CCTTGGATCTTGGCGGTTCT R: CACTGCCATGCATGATGCTC	NM_001206359
PGK1	F: CCTTGGATCTTGGCGGTTCT R: CACTGCCATGCATGATGCTC	NM_001099932
PDK1	F: CGTGCTGGGAATCAGCAAAC R: GCTCGAAGTCCGTCTCCTTC	NM_001159608
LDHA	F: ATCCTGTGGACGGAAGCATT R: AGGTGATAACAGTGGGTGCG	NM_001172363
COX1	F: GGAGGTCTAACGGGCATTGT R: ACCCGGAGAATAGGGGAAT	NP_008636
COX2	F: CCAAGACGCCACTTCACCCATC R: TGGGCATCCATTGTGCTAGTGTG	NP_008637.1
COX3	F: AAGACGCCACTTCACCCATC R: TCTTGGGCATCCATTGTGCT	NP_008640
ATP6	F: CCGCACAATCTCGATCCAAC R: AGTTGTGTGGTGGGTGTGAA	NP_008639
ATP8	F: GCCACAAC TAGATCATCCACATG R: GATTCTGGGCTTGCTGGGTATG	NP_008638.1

Table 2. qPCR reaction system.

Cycle Step	Repetition Number	Temperature	Times
premutability	1	95 °C	30 s
cyclic response	40	95 °C	10 s
		60 °C	30 s
		95 °C	15 s
		60 °C	60 s
Dissolution curves	1	95 °C	15 s

2.4. TMRE Assay of Mitochondrial Membrane Potential (MMP)

The culture medium of PFFs treated with CPI for 48 h was aspirated, cells were washed with PBS, and a working solution was prepared in accordance with TMRE: serum-free medium = 1:1000 using the TMRE kit (Shanghai Beyotime Biotech, Shanghai, China, Item No. C2001S). Subsequently, 1 mL of the working solution was added to each well and incubated for 45 min in a 37 °C incubator, shielded from light. At the end of the incubation period, the supernatant was aspirated, and the cells were washed twice with a pre-warmed cell culture solution. Following this, 500 µL of the pre-warmed cell culture solution containing serum and phenol red was added. The observation and documentation of cell fluorescence were conducted under a fluorescence microscope, and images were quantified using ImageJ (13.0.6).

2.5. Measurement of Lactic Acid Content

A lactic acid content kit purchased from Shanghai Biyuntian Biotechnology Co., Ltd., Shanghai, China, (Item No.: C0017) was used for measuring lactic acid content. The extraction process involved a cell-to-extract ratio of 500~1000:1. Specifically, 1 mL of extract was added to the designated volume of cells, and the cells were subjected to ultrasonic ice bath treatment for cell lysis at a power of 300 W, with ultrasonication intervals of 3 s and intervals of 7 s, totaling 3 min. The process was conducted at 4 °C, followed by centrifugation at 12,000× g for 10 min. Subsequently, 0.8 mL of the supernatant was collected, and 0.15 mL of extraction solution II was slowly added with careful blowing and mixing until no gas bubbles were observed. The mixture underwent a 4 °C, 12,000× g centrifugation for an additional 10 min, after which the supernatant was collected. The lactic acid content was then determined using the lactic acid content kit by measuring absorbance at 570 wavelengths, and the data were recorded. The lactic acid content was calculated using a standard curve.

2.6. Somatic Cell Nuclear Transfer

The CPI-treated PFFs were subjected to digestion and resuspension with DPBS-PVA. In vitro matured oocytes were enucleated through blind aspiration. Blind aspiration is performed with a microfine glass tube under the first polar body to aspirate the first polar body and chromosomes in mid-division and some of the surrounding cytoplasm to ensure removal of the genetic material. Single donor cells were aspirated into the microinjection needle and injected into the zona pellucida of the nucleated oocytes. The reconstructed embryos were then equilibrated in PZM-3 for 1 h. Subsequently, the embryos underwent transfer into an electrofusion solution, and fusion was induced by two direct current (DC) pulses at 100 v/mm for 100 µs. After fusion, they were transferred back into PZM-3 and placed in the incubator. Cleavage rates were observed on the second day of incubation, while blastocysts were collected on the seventh day to record embryonic development efficiency and the number of blastocyst cells.

2.7. Transcriptome Sequencing Analysis

Standard transcriptome sequencing was conducted on PFFs subjected to a 50 µM CPI treatment and untreated NC, with four biological replicates for each condition. The treated cells were denoted as CPI 1–4, while the untreated ones were labeled as NC1–4.

Total cellular RNA was extracted by the Trizol method as the starting RNA for library construction, and cDNA was synthesized for PCR amplification to construct the library. After checking the quality of the library, Illumina sequencing was performed to obtain clean reads, and HISAT2 (v2.0.5) was used to compare the clean reads with the positional information on the reference genome, count the number of reads for each gene, and calculate the FPKM value. DESeq2 software (1.20.0) was used to perform the statistical analysis to select differentially expressed genes. $\log_2(\text{Foldchange}) > 1$ & p value < 0.05 . The cluster Profiler software (3.8.1) was used to map the DEGs to GO and KEGG databases, and the differential gene sets were analyzed for GO functional annotation and KEGG pathway enrichment. All the above work was conducted by Beijing Novozymes Technology Co., Beijing, China.

2.8. Statistical Analysis

The data analysis was conducted using SPSS 20.0, employing t -tests for pairwise comparisons between groups, and one-way ANOVA for analyzing differences among two or more groups.

3. Results

3.1. CPI Treatment Downregulated Oxidative Phosphorylation Gene Expression but Upregulated Glycolytic Enzyme Gene Expression in PFFs

PFFs were treated with different concentrations of CPI (10 μM , 50 μM , and 100 μM) for 48 h. The mRNA expression levels of three oxidative phosphorylation genes (Mitochondrially Encoded Cytochrome C Oxidase I (COX1), Mitochondrially Encoded Cytochrome C Oxidase III (COX3), Mitochondrially Encoded ATP Synthase Membrane Subunit 6 (ATP6)) and three glycolysis genes (Phosphoglycerate Kinase 1 (PGK1), Pyruvate Dehydrogenase Kinase 1 (PDK1), Lactate Dehydrogenase A (LDHA)) in the treated PFFs were measured by qPCR. As depicted in Figure 1, the 10 μM CPI incubation had no significant effect on the expression of all tested genes. The addition of 50 μM CPI to the PFF culture medium significantly decreased ATP6 and increased PGK1 mRNA abundance, as compared to the NC group. The supplementation of 100 μM CPI to the culture medium of PFFs significantly downregulated the transcript level of oxidative phosphorylation genes COX3 and ATP6 and upregulated the transcript level of glycolytic enzyme gene PGK1. These results indicate that treating PFFs with CPI at a concentration of 50 or 100 μM inhibits the expression of oxidative phosphorylation-related genes while promoting the expression of a glycolytic gene, PGK1.

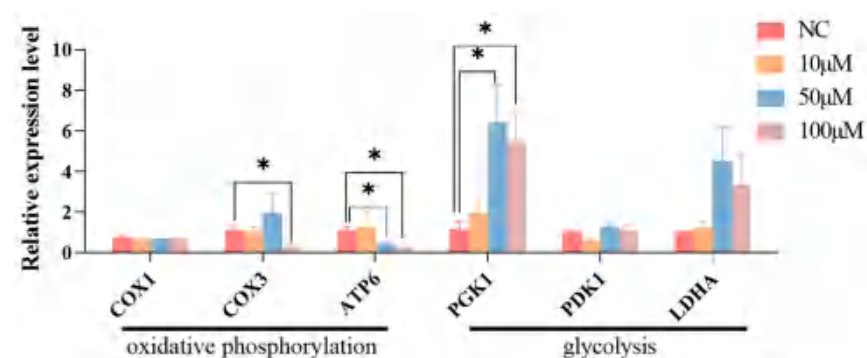


Figure 1. Impacts of CPI treatment on oxidative phosphorylation and glycolysis-related gene expression in PFFs. Data are presented as “Means \pm S.E.”. Three biological replicates in each group. “*” indicates a significant difference between two groups in the same column ($p < 0.05$).

3.2. CPI Treatment of PFFs Resulted in Downregulation of MMP and Upregulation of Lactate Content

Mitochondria provide energy for the host cells by producing ATPs, predominantly through oxidative phosphorylation. The MMP (Mn) reflects oxidative phosphorylation

status of the cells. Therefore, we investigated the effects of CPI treatment on the MMP of PFFs. The results revealed a significant decrease in cellular MMP following the 50 μ M and 100 μ M CPI treatment, compared to the NC group (Figure 2A,B). To explore the effect of CPI on the PFF glycolytic pathway, we also assessed the lactate content, which is an index of glycolysis, of PFFs treated with 50 μ M CPI. The results demonstrated that lactate content in the CPI-treated group was significantly increased, by one-fold, compared to the NC group (Figure 2C). This observation indicates that CPI suppresses the cellular oxidative phosphorylation pathway but promotes the glycolytic pathway, resulting in a drop in cellular MMP and an upregulation of lactate content.

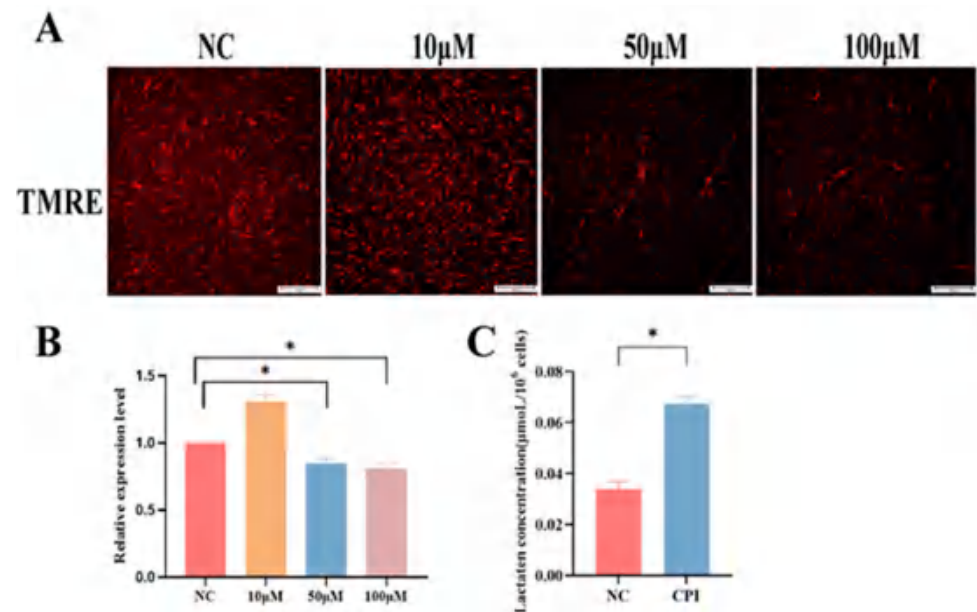


Figure 2. Impacts of CPI treatment on MMP and lactate content in PFFs: (A) effects of CPI treatment on MMP of PFFs, PFFs (red) were stained by TMRE. NC represents the negative NC group, PFFs were treated with 10 μ M, 50 μ M, 100 μ M CPI; Original magnification: $\times 100$, bar: 100 μ m; (B) the relative fluorescence intensity of TMRE probe; and (C) effects of 50 μ M CPI treatment on lactic acid content in PFFs. Data are presented as “Means \pm S.E.”. Three biological replicates in each group. “*” indicates a significant difference between two groups in the same column ($p < 0.05$).

3.3. CPI Treatment of Donor Cells PFFs Significantly Enhanced the Developmental Competence of Cloned Embryos

To investigate the influence of CPI treatment of donor cells on subsequent development of porcine cloned embryos, we compared the developmental indexes (cleavage rate, blastocyst rate, and the number of blastocyst cells) of cloned embryos constructed from 50 μ M CPI-treated PFFs and NC PFFs. The results revealed that the blastocyst rate ($49.33 \pm 3.52\%$ vs. $38.00 \pm 2.0\%$, $p < 0.05$) in the treated group was significantly higher than that in the NC group (Table 3). These findings suggest that incubating donor cells with the oxidative phosphorylation inhibitor CPI promotes the subsequent development of cloned pig embryos.

Table 3. Effects of CPI treatment of donor cells on subsequent development of porcine cloned embryos.

Groups	Cleavage Rate (n)	Four-Cell Stage Rate (n)	Blastocyst Rate (n)	Number of Blastocyst Cells Mean
NC	$67.00 \pm 1.91\%$ (67/100)	$37.00 \pm 1.91\%$ (37/100)	$38.00 \pm 2.0\%$ (38/100)	34.13 ± 13.38
CPI	$69.33 \pm 1.15\%$ (52/75)	$38.67 \pm 1.15\%$ (40/75)	$49.33 \pm 3.52\%$ (43/75) *	43.19 ± 18.52

Data is presented as “Means \pm S.E.”. “*” indicates a significant difference between two groups in the same column ($p < 0.05$).

3.4. Transcriptomic Analysis of PFFs Treated with CPI

To investigate the impacts of donor cell energy metabolism pathways on the development of porcine cloned embryos, PFFs without any treatment and those treated with 50 μ M CPI for 48 h were collected for transcriptome sequencing. Based on the principal component analysis (PCA) results, the NC- and CPI-treated groups exhibited clear grouping, demonstrating significant inter-group differences and intra-group reproducibility (Figure 3A). A total of 10,133 differentially expressed genes (DEGs) were identified between the two groups, comprising 5007 upregulated genes and 5126 downregulated genes in the CPI-treated group (Figure 3B). Gene Ontology (G.O.) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were conducted on the DEGs identified through RNA sequencing. The G.O. analysis revealed enrichment of DEGs in metabolism-related pathways such as the mitochondrial inner membrane protein complex, respirator, inner mitochondrial membrane respiratory chain complex, NADH dehydrogenase activity, oxidoreductase activity and other pathways (Figure 3C). KEGG analysis demonstrated that DEGs between the two cell groups were also enriched in several metabolism-related pathways, including reactive oxygen species, oxidative phosphorylation, thermogenesis, and other pathways. (Figure 3D).

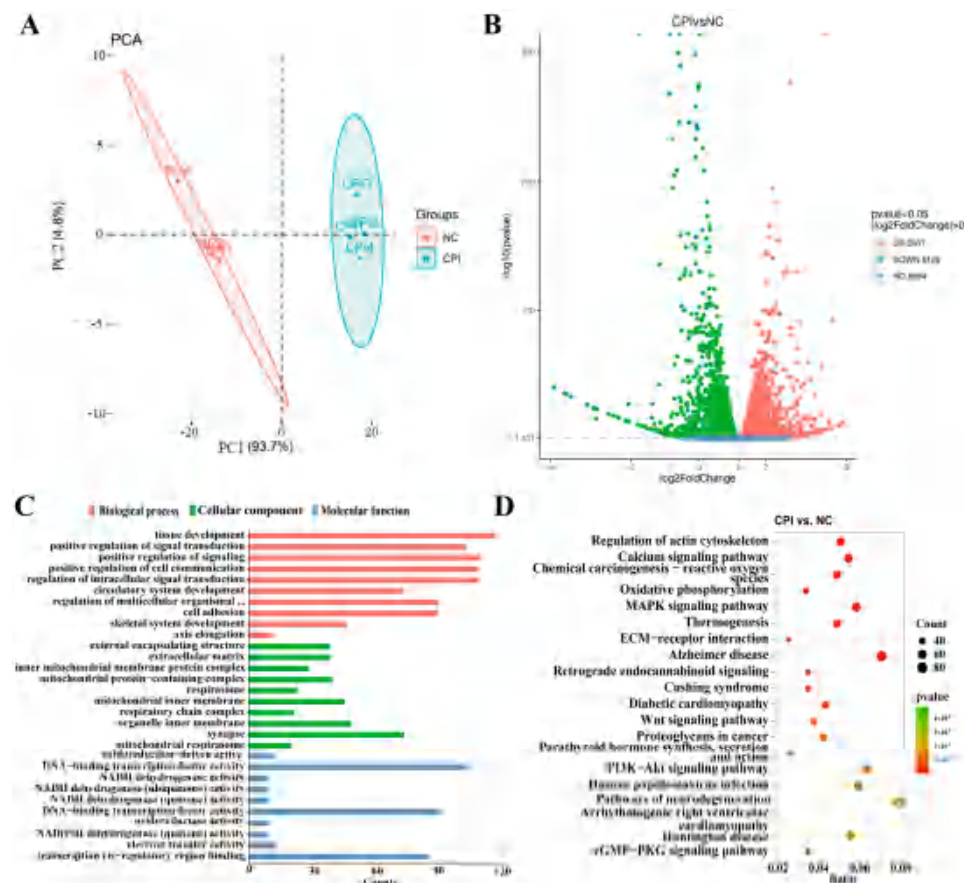


Figure 3. Transcriptome sequencing analysis of CPI-Treated Group and NC Group: (A) principal component analysis (PCA) of both groups of cells; (B) volcano plot of differentially expressed genes (DEGs) between two groups of cells; (C) functional enrichment analysis of DEGs based on the Gene Ontology (G.O.) database, showcasing the top 10 pathways in each subcategory; and (D) functional enrichment analysis of DEGs based on the Kyoto Encyclopedia of Genes and Genomes (KEGG), presenting the top 20 KEGG pathways.

3.5. Interaction Analysis of Oxidative Phosphorylation-Related Pathway Networks

The ROS signaling pathway, oxidative phosphorylation pathway, and thermogenic signaling pathway, which are closely related to energy metabolism in the KEGG enrichment

analysis, were mapped for network interactions (Figure 4A). The results showed that the key node genes, including cytochrome genes (Mitochondrially Encoded Cytochrome C Oxidase I: *COX1*, Mitochondrially Encoded Cytochrome C Oxidase II: *COX2*, Mitochondrially Encoded Cytochrome C Oxidase III: *COX3*) and ATP synthase genes (Mitochondrially Encoded ATP Synthase Membrane Subunit 6: *ATP6*, Mitochondrially Encoded ATP Synthase Membrane Subunit 8: *ATP8*) had a significantly lower expression in the CPI-treated group than in the NC group (Figure 4B). Cytochrome genes and ATP synthase genes have been shown to inhibit the oxidative phosphorylation pathway [32,33]. The downregulated expression of these two families of genes in the CPI-treated cells indicated that CPI treatment significantly suppressed the oxidative phosphorylation pathways in the cells, further confirming that CPI is an effective oxidative phosphorylation inhibitor. To validate the expression level of five node genes, including *COX1*, *COX2*, *COX3*, *ATP6* and *ATP8*, their mRNA expression levels in 50 μ M CPI-treated and NC PPFs were measured by qPCR. As shown in Figure 4C, the expression levels of the tested five genes were lower in the CPI-treated group than in the control group, which was consistent with the downregulation trend in the RNA sequencing results, indicating that the RNA sequencing data were accurate and reliable.

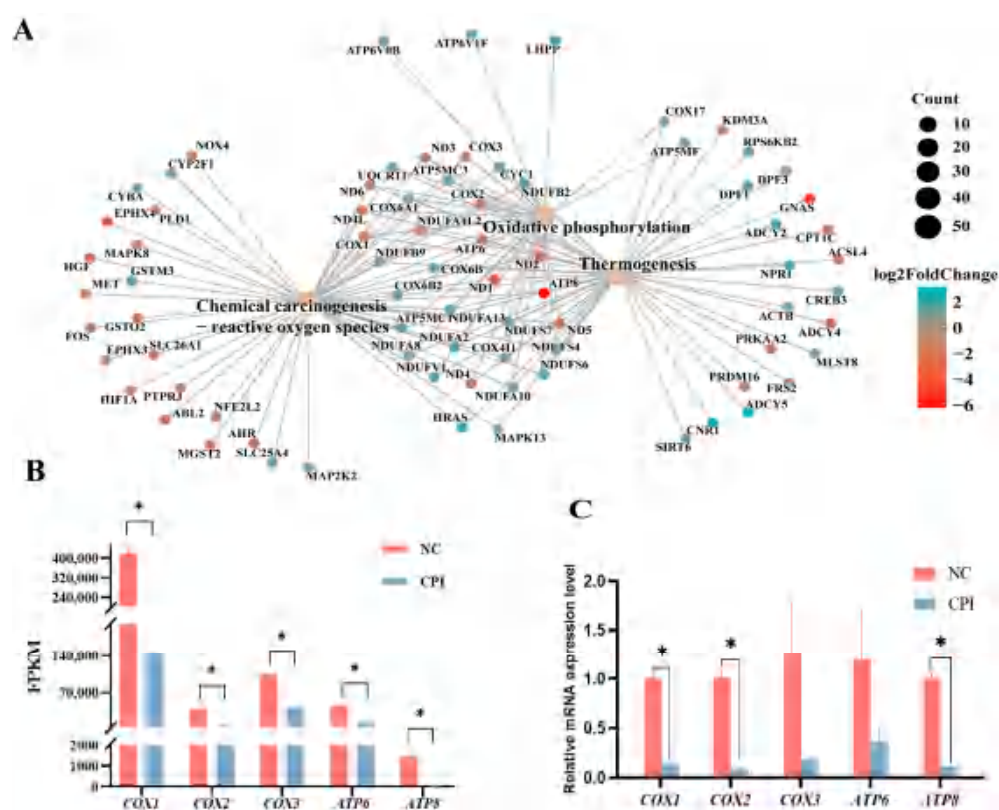


Figure 4. Interaction analysis of oxidative phosphorylation-related pathway networks: (A) interaction map of oxidative phosphorylation, ROS, and thermogenic signaling pathway networks; (B) comparison of FPKM values of node genes between CPI-Treated Group and NC Group; and (C) qPCR validation of RNA sequencing-identified DEGs. Data are presented as “Means \pm S.E.”. Three biological replicates in each group. “*” indicates a significant difference between two groups in the same column ($p < 0.05$).

4. Discussion

CPI, as an analogue of alpha-lipoic acid, suppresses oxidative phosphorylation by inhibiting the expression of pyruvate dehydrogenase (*PDH*), which is an enzyme complex within the mitochondria pivotal that catalyzes the conversion of pyruvate to acetyl CoA [34–36]. The interruption of *PDH* expression by CPI halts the entry of pyruvate into

the tricarboxylic acid cycle, disrupting oxidative phosphorylation and diminishing ATP production [37]. Furthermore, the CPI-mediated PDH blockade triggers the accumulation of acetyl-CoA and promotes the glycolytic pathway to meet metabolic requirements [38]. Our qPCR results revealed that the 10 μ M CPI treatment exerted no significant effect on the gene expression of energy metabolism pathways in donor cells, whereas the 50 μ M and 100 μ M CPI supplementation significantly inhibited the expression of oxidative phosphorylation-related genes. Similar results were found in previous research, which demonstrated that a 10 μ M CPI treatment failed to stimulate the PFF oxidative phosphorylation pathway, while a higher concentration of 100 μ M CPI inhibited oxidative phosphorylation [39].

Inhibition of the oxidative phosphorylation pathway disrupts the electron transport chain's function in the mitochondria, causing a reduction in the potential gradient between the mitochondrial membranes [40]. In this study, MMP downregulation was observed in 50 μ M and 100 μ M CPI-treated PFFs. Previous research has shown that 50 μ M and 100 μ M CPI can decrease the MMP of PFFs [41]. The reprogramming of mouse ear fibroblasts (MEFs) to induce pluripotent stem cells (iPSCs) also involves MMP reduction [42].

During glycolysis, glucose is decomposed into pyruvate, which is then metabolized into lactate in the cytoplasm; therefore, the level of lactate is positively correlated with the intensity of intracellular glycolysis [43]. It has been shown that upregulation of the glycolytic pathway is accompanied by an increase in lactate content [44–46]. The significant elevation of lactate levels in the CPI-treated PFFs strongly suggests that the glycolysis process is enhanced in PFFs following CPI treatment.

Among the pathways analyzed by KEGG enrichment in this study, the Wnt, AMPK, and PI3K/AKT signaling pathways are associated with energy metabolism, in addition to the ROS signaling pathway [47], oxidative phosphorylation pathway [45], and thermogenic signaling pathway [48]. It has been shown that activation of the Wnt signaling pathway promotes upregulation of ATP synthase, leading to increased intra-cellular lactate secretion and upregulation of the glycolytic pathway [49]. AMPK enhances the glycolytic pathway by upregulating the expression of glycolytic enzymes [50]. The upregulation of the PI3K/AKT pathway promotes the activities of glycolytic pyruvate kinase and hexokinase, which increase ATP production and promote glycolysis and lactate production in breast cancer cells [51].

In the present study, the blastocyst rate of SCNT embryos cloned from CPI-treated PFFs was significantly higher compared to the NC group treatment of PFFs with 100 μ M CoCl_2 (a glycolysis inducer) and also resulted in a significantly higher blastocyst rate and blastocyst cell number in subsequent SCNT embryos, compared to NC [31]. Similarly, treatment of PFFs with 10 μ M PS-48, another glycolysis inducer, led to significantly higher cleavage and blastocyst rates in subsequent cloned embryos compared to NC [30]. These drug treatments all enhance the developmental efficiency of SCNT embryos by altering energy metabolic pathways in donor cells. However, it is important to note that the blastocyst rate and blastocyst cell number may not fully represent the uterine survival rate of cloned embryos [52]. Whether the CPI treatment of donor cells can improve the full-term developmental competence of subsequent cloned embryos remains unknown and requires further investigation.

Our results showed that CPI treatment increased the lactic acid level in donor cells. Lactate has been shown to be essential for the early development of mammalian embryos, and deprivation of lactate leads to significant deletion of H3K18lac and to the failure of major ZAG activation in mouse 2-cell embryos [53]. It has also been shown that lactate promotes epigenetic reprogramming through histone H3K27 acetylation [54]. Increased levels of histone acetylation promote transcriptional activation and increased expression levels of developmentally relevant genes in early SCNT embryos, ultimately enhancing the developmental potential of SCNT embryos [55]. Therefore, CPI treatment improves the developmental efficiency of subsequent cloned embryos.

5. Conclusions

Treatment of PFFs with the oxidative phosphorylation inhibitor CPI effectively shifts the cellular energy metabolic pathways from oxidative phosphorylation to glycolysis and enhances the developmental potential of subsequent SCNT embryos. This study provides a simple new method to improve the efficiency of pig cloning, which will be beneficial for promoting the development and application of pig SCNT technology.

Author Contributions: Conceptualization, J.C. and Z.L. (Zicong Li); methodology, J.C. and Y.D.; software, Y.D. and J.C.; validation, J.C., Y.D. and Z.L. (Zheng Li); formal analysis, S.W. and J.C.; investigation, S.W. and J.C.; resources, Y.D.; data curation, J.C. and Y.D.; writing—original draft preparation, J.C., Y.D. and Z.L. (Zheng Li); writing—review and editing, Z.L. (Zicong Li) and E.Z.; project administration, Z.L. (Zicong Li), E.Z. and Z.W.; funding acquisition, Z.L. (Zicong Li), E.Z. and Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by one grant from the Department of Agriculture and Rural Affairs of Guangdong Province, China (grant number: 2022-XPY-00-005), and a grant from the Department of Science and Technology of Guangdong Province, China (grant number: 2018B030313011).

Institutional Review Board Statement: The animal study protocol was approved by the Ethics Committee of the Experimental Animal Center of South China Agricultural University (License No: SYXK-2019-0136).

Informed Consent Statement: Informed consent was not applicable as there were no human subjects involved in this article.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Yadav, P.S.; Kumar, D.; Saini, M.; Sharma, R.K.; Dua, S.; Selokar, N.L.; Bansal, S.; Punetha, M.; Gupta, A.; Kumar, R.; et al. Evaluation of postnatal growth, hematology, telomere length and semen attributes of multiple clones and re-clone of superior buffalo breeding bulls. *Theriogenology* **2024**, *213*, 24–33. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Shi, J.; Xiao, L.; Tan, B.; Luo, L.; Li, Z.; Hong, L.; Yang, J.; Cai, G.; Zheng, E.; Wu, Z.; et al. Comparative evaluation of production performances of cloned pigs derived from superior duroc boars. *Anim. Reprod. Sci.* **2022**, *244*, 107049. [\[CrossRef\]](#)
3. Shi, J.; Tan, B.; Luo, L.; Li, Z.; Hong, L.; Yang, J.; Cai, G.; Zheng, E.; Wu, Z.; Gu, T. Assessment of the growth and reproductive performance of cloned pietrain boars. *Animals* **2020**, *10*, 2053. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Loi, P.; Ptak, G.; Barboni, B.; Fulka, J.J.; Cappai, P.; Clinton, M. Genetic rescue of an endangered mammal by cross-species nuclear transfer using post-mortem somatic cells. *Nat. Biotechnol.* **2001**, *19*, 962–964. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Gomez, M.C.; Pope, C.E.; Giraldo, A.; Lyons, L.A.; Harris, R.F.; King, A.L.; Cole, A.; Godke, R.A.; Dresser, B.L. Birth of african wildcat cloned kittens born from domestic cats. *Cloning Stem Cells* **2004**, *6*, 247–258. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Huang, L.; Hua, Z.; Xiao, H.; Cheng, Y.; Xu, K.; Gao, Q.; Xia, Y.; Liu, Y.; Zhang, X.; Zheng, X.; et al. Crispr/cas9-mediated *ApoE*^{−/−} and *LDLR*^{−/−} double gene knockout in pigs elevates serum ldl-c and tc levels. *Oncotarget* **2017**, *8*, 37751–37760. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Huang, J.; Guo, X.; Fan, N.; Song, J.; Zhao, B.; Ouyang, Z.; Liu, Z.; Zhao, Y.; Yan, Q.; Yi, X.; et al. Rag1/2 knockout pigs with severe combined immunodeficiency. *J. Immunol.* **2014**, *193*, 1496–1503. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Langin, M.; Reichart, B.; Steen, S.; Sjoberg, T.; Paskevicius, A.; Liao, Q.; Qin, G.; Mokolke, M.; Mayr, T.; Radan, J.; et al. Cold non-ischemic heart preservation with continuous perfusion prevents early graft failure in orthotopic pig-to-baboon xenotransplantation. *Xenotransplantation* **2021**, *28*, e12636. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Rosales, I.A.; Kinoshita, K.; Maenaka, A.; Anne, L.H.I.; Selig, M.K.; Laguerre, C.M.; Collins, A.B.; Ayares, D.; Cooper, D.; Colvin, R.B. De novo membranous nephropathy in a pig-to-baboon kidney xenograft: A new xenograft glomerulopathy. *Am. J. Transplant.* **2024**, *24*, 30–36. [\[CrossRef\]](#)
10. Zeng, F.; Liao, S.; Kuang, Z.; Zhu, Q.; Wei, H.; Shi, J.; Zheng, E.; Xu, Z.; Huang, S.; Hong, L.; et al. Genetically engineered pigs as efficient salivary gland bioreactors for production of therapeutically valuable human nerve growth factor. *Cells* **2022**, *11*, 2378. [\[CrossRef\]](#)
11. Wang, M.; Sun, Z.; Yu, T.; Ding, F.; Li, L.; Wang, X.; Fu, M.; Wang, H.; Huang, J.; Li, N.; et al. Large-scale production of recombinant human lactoferrin from high-expression, marker-free transgenic cloned cows. *Sci. Rep.* **2017**, *7*, 10733. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Tabar, V.; Tomishima, M.; Panagiotakos, G.; Wakayama, S.; Menon, J.; Chan, B.; Mizutani, E.; Al-Shamy, G.; Ohta, H.; Wakayama, T.; et al. Therapeutic cloning in individual parkinsonian mice. *Nat. Med.* **2008**, *14*, 379–381. [\[CrossRef\]](#) [\[PubMed\]](#)

13. Sui, L.; Danzl, N.; Campbell, S.R.; Viola, R.; Williams, D.; Xing, Y.; Wang, Y.; Phillips, N.; Poffenberger, G.; Johannesson, B.; et al. Beta-cell replacement in mice using human type 1 diabetes nuclear transfer embryonic stem cells. *Diabetes* **2018**, *67*, 26–35. [\[CrossRef\]](#)
14. Wang, X.; Shi, J.; Cai, G.; Zheng, E.; Liu, D.; Wu, Z.; Li, Z. Overexpression of mbd3 improves reprogramming of cloned pig embryos. *Cell. Reprogramm.* **2019**, *21*, 221–228. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Zhao, H.; Xie, S.; Zhang, N.; Ao, Z.; Wu, X.; Yang, L.; Shi, J.; Mai, R.; Zheng, E.; Cai, G.; et al. Source and follicular fluid treatment during the in vitro maturation of recipient oocytes affects the development of cloned pig embryo. *Cell. Reprogramm.* **2020**, *22*, 71–81. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Dong, Y.; Wu, X.; Peng, X.; Yang, L.; Tan, B.; Zhao, H.; Zheng, E.; Hong, L.; Cai, G.; Wu, Z.; et al. Knockdown of yy1 inhibits xist expression and enhances cloned pig embryo development. *Int. J. Mol. Sci.* **2022**, *23*, 14572. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Costa, G.R.; Souza, R.E.; Zago, F.C.; Aguiar, L.H.; Rodriguez-Villamil, P.; Ongaratto, F.L.; Ambrosio, C.E.; Miglino, M.A.; Rodrigues, J.L.; Forell, F.; et al. Effects of fusion-activation interval and embryo aggregation on in vitro and in vivo development of bovine cloned embryos. *Res. Vet. Sci.* **2019**, *123*, 91–98. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Meng, F.; Stamms, K.; Bennewitz, R.; Green, A.; Obach, F.; Turner, P.; Wei, J.; Obach, B. Targeted histone demethylation improves somatic cell reprogramming into cloned blastocysts but not postimplantation bovine conceptidagger. *Biol. Reprod.* **2020**, *103*, 114–125. [\[CrossRef\]](#)
19. Qu, P.; Qing, S.; Liu, R.; Qin, H.; Wang, W.; Qiao, F.; Ge, H.; Liu, J.; Zhang, Y.; Cui, W.; et al. Effects of embryo-derived exosomes on the development of bovine cloned embryos. *PLoS ONE* **2017**, *12*, e174535. [\[CrossRef\]](#)
20. Callesen, H.; Liu, Y.; Pedersen, H.S.; Li, R.; Schmidt, M. Increasing efficiency in production of cloned piglets. *Cell. Reprogramm.* **2014**, *16*, 407–410. [\[CrossRef\]](#)
21. Wells, D.N.; Laible, G.; Tucker, F.C.; Miller, A.L.; Oliver, J.E.; Xiang, T.; Forsyth, J.T.; Berg, M.C.; Cockrem, K.; L’Huillier, P.J.; et al. Coordination between donor cell type and cell cycle stage improves nuclear cloning efficiency in cattle. *Theriogenology* **2003**, *59*, 45–59. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Christofk, H.R.; Vander, H.M.; Harris, M.H.; Ramanathan, A.; Gerszten, R.E.; Wei, R.; Fleming, M.D.; Schreiber, S.L.; Cantley, L.C. The m2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature* **2008**, *452*, 230–233. [\[CrossRef\]](#)
23. El-Sanea, A.M.; Abdoon, A.; Kandil, O.M.; El-Toukhy, N.E.; El-Maaty, A.; Ahmed, H.H. Effect of oxygen tension and antioxidants on the developmental competence of buffalo oocytes cultured in vitro. *Vet. World* **2021**, *14*, 78–84. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Rybouchkin, A.; Kato, Y.; Tsunoda, Y. Role of histone acetylation in reprogramming of somatic nuclei following nuclear transfer. *Biol. Reprod.* **2006**, *74*, 1083–1089. [\[CrossRef\]](#)
25. Karagiannis, P.; Takahashi, K.; Saito, M.; Yoshida, Y.; Okita, K.; Watanabe, A.; Inoue, H.; Yamashita, J.K.; Todani, M.; Nakagawa, M.; et al. Induced pluripotent stem cells and their use in human models of disease and development. *Physiol. Rev.* **2019**, *99*, 79–114. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Prigione, A.; Rohwer, N.; Hoffmann, S.; Mlody, B.; Drews, K.; Bukowiecki, R.; Blumlein, K.; Wanker, E.E.; Ralser, M.; Cramer, T.; et al. Hif1alpha modulates cell fate reprogramming through early glycolytic shift and upregulation of pdk1-3 and pkm2. *Stem. Cells.* **2014**, *32*, 364–376. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Prigione, A.; Lichtner, B.; Kuhl, H.; Struys, E.A.; Wamelink, M.; Lehrach, H.; Ralser, M.; Timmermann, B.; Adjaye, J. Human induced pluripotent stem cells harbor homoplasmic and heteroplasmic mitochondrial DNA mutations while maintaining human embryonic stem cell-like metabolic reprogramming. *Stem Cells* **2011**, *29*, 1338–1348. [\[CrossRef\]](#)
28. Shakya, A.; Cooksey, R.; Cox, J.E.; Wang, V.; McClain, D.A.; Tantin, D. Oct1 loss of function induces a coordinate metabolic shift that opposes tumorigenicity. *Nat. Cell Biol.* **2009**, *11*, 320–327. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Zhu, S.; Li, W.; Zhou, H.; Wei, W.; Ambasudhan, R.; Lin, T.; Kim, J.; Zhang, K.; Ding, S. Reprogramming of human primary somatic cells by oct4 and chemical compounds. *Cell Stem Cell* **2010**, *7*, 651–655. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Luo, C.; Wang, Z.; Wang, J.; Yun, F.; Lu, F.; Fu, J.; Liu, Q.; Shi, D. Individual variation in buffalo somatic cell cloning efficiency is related to glycolytic metabolism. *Sci. China Life Sci.* **2022**, *65*, 2076–2092. [\[CrossRef\]](#)
31. Cecil, R.F.; Chen, P.R.; Benne, J.A.; Hord, T.K.; Spate, L.D.; Samuel, M.S.; Prather, R.S. Chemical simulation of hypoxia in donor cells improves development of somatic cell nuclear transfer-derived embryos and increases abundance of transcripts related to glycolysis. *Mol. Reprod. Dev.* **2020**, *87*, 763–772. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Mordhorst, B.R.; Murphy, S.L.; Ross, R.M.; Benne, J.A.; Samuel, M.S.; Cecil, R.F.; Redel, B.K.; Spate, L.D.; Murphy, C.N.; Wells, K.D.; et al. Pharmacologic treatment of donor cells induced to have a arburg effect-like metabolism does not alter embryonic development in vitro or survival during early gestation when used in somatic cell nuclear transfer in pigs. *Mol. Reprod. Dev.* **2018**, *85*, 290–302. [\[CrossRef\]](#) [\[PubMed\]](#)
33. 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. *Methods* **2001**, *25*, 402–408. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Guan, F.; Yu, B.; Qi, G.X.; Luo, J. A study on the effect of sodium azide on myocardial cell viability. *J. China Med. Univ.* **2006**, 345–347.
35. Wu, L.; Zhao, J.; Cao, K.; Liu, X.; Cai, H.; Wang, J.; Li, W.; Chen, Z. Oxidative phosphorylation activation is an important characteristic of dox resistance in hepatocellular carcinoma cells. *Cell Commun. Signal.* **2018**, *16*, 6. [\[CrossRef\]](#)
36. Sturmey, R.G.; Leese, H.J. Energy metabolism in pig oocytes and early embryos. *Reproduction* **2003**, *126*, 197–204. [\[CrossRef\]](#)

37. Pardee, T.S.; Lee, K.; Luddy, J.; Maturo, C.; Rodriguez, R.; Isom, S.; Miller, L.D.; Stadelman, K.M.; Levitan, D.; Hurd, D.; et al. A phase I study of the first-in-class antimitochondrial metabolism agent, cpi-613, in patients with advanced hematologic malignancies. *Clin. Cancer Res.* **2014**, *20*, 5255–5264. [[CrossRef](#)] [[PubMed](#)]
38. Reddy, V.B.; Boteju, L.; Boteju, A.; Shen, L.; Kassahun, K.; Reddy, N.; Sheldon, A.; Luther, S.; Hu, K. In vitro and in vivo metabolism of a novel antimitochondrial cancer metabolism agent, cpi-613, in rat and human. *Drug. Metab. Dispos.* **2022**, *50*, 361–373. [[CrossRef](#)]
39. Mordhorst, B.R.; Kerns, K.C.; Schauflinger, M.; Zigo, M.; Murphy, S.L.; Ross, R.M.; Wells, K.D.; Green, J.A.; Sutovsky, P.; Prather, R.S. Pharmacologic treatment with cpi-613 and ps48 decreases mitochondrial membrane potential and increases quantity of autolysosomes in porcine fibroblasts. *Sci. Rep.* **2019**, *9*, 9417. [[CrossRef](#)]
40. Zhao, Z.; Mei, Y.; Wang, Z.; He, W. The effect of oxidative phosphorylation on cancer drug resistance. *Cancers* **2022**, *15*, 62. [[CrossRef](#)]
41. Mordhorst, B.R.; Murphy, S.L.; Schauflinger, M.; Rojas, S.S.; Ji, T.; Behura, S.K.; Wells, K.D.; Green, J.A.; Prather, R.S. Porcine fetal-derived fibroblasts alter gene expression and mitochondria to compensate for hypoxic stress during culture. *Cell. Reprogramm.* **2018**, *20*, 225–235. [[CrossRef](#)]
42. Cui, G.; Zhou, J.; Sun, J.; Kou, X.; Su, Z.; Xu, Y.; Liu, T.; Sun, L.; Li, W.; Wu, X.; et al. Wd repeat domain 82 (wdr82) facilitates mouse ipscs generation by interfering mitochondrial oxidative phosphorylation and glycolysis. *Cell. Mol. Life Sci.* **2023**, *80*, 218. [[CrossRef](#)] [[PubMed](#)]
43. Nishioku, T.; Anzai, R.; Hiramatsu, S.; Terazono, A.; Nakao, M.; Moriyama, M. Lactate dehydrogenase a inhibition prevents rankl-induced osteoclastogenesis by reducing enhanced glycolysis. *J. Pharmacol. Sci.* **2023**, *153*, 197–207. [[CrossRef](#)] [[PubMed](#)]
44. Yonashiro, R.; Eguchi, K.; Wake, M.; Takeda, N.; Nakayama, K. Pyruvate dehydrogenase pdh-e1beta controls tumor progression by altering the metabolic status of cancer cells. *Cancer Res.* **2018**, *78*, 1592–1603. [[CrossRef](#)] [[PubMed](#)]
45. Matoba, S.; Kang, J.G.; Patino, W.D.; Wragg, A.; Boehm, M.; Gavrilova, O.; Hurley, P.J.; Bunz, F.; Hwang, P.M. P53 regulates mitochondrial respiration. *Science* **2006**, *312*, 1650–1653. [[CrossRef](#)] [[PubMed](#)]
46. Rho, H.; Terry, A.R.; Chronis, C.; Hay, N. Hexokinase 2-mediated gene expression via histone lactylation is required for hepatic stellate cell activation and liver fibrosis. *Cell Metab.* **2023**, *35*, 1406–1423. [[CrossRef](#)] [[PubMed](#)]
47. Nolfi-Donagan, D.; Braganza, A.; Shiva, S. Mitochondrial electron transport chain: Oxidative phosphorylation, oxidant production, and methods of measurement. *Redox Biol.* **2020**, *37*, 101674. [[CrossRef](#)] [[PubMed](#)]
48. Zorov, D.B.; Andrianova, N.V.; Babenko, V.A.; Bakeeva, L.E.; Zorov, S.D.; Zorova, L.D.; Pevsner, I.B.; Popkov, V.A.; Plotnikov, E.Y.; Silachev, D.N. Nonphosphorylating oxidation in mitochondria and related processes. *Biochemistry* **2020**, *85*, 1570–1577. [[CrossRef](#)] [[PubMed](#)]
49. Mo, Y.; Wang, Y.; Zhang, L.; Yang, L.; Zhou, M.; Li, X.; Li, Y.; Li, G.; Zeng, Z.; Xiong, W.; et al. The role of wnt signaling pathway in tumor metabolic reprogramming. *J. Cancer* **2019**, *10*, 3789–3797. [[CrossRef](#)]
50. Marsin, A.S.; Bertrand, L.; Rider, M.H.; Deprez, J.; Beauloye, C.; Vincent, M.F.; Van den Berghe, G.; Carling, D.; Hue, L. Phosphorylation and activation of heart pfk-2 by ampk has a role in the stimulation of glycolysis during ischaemia. *Curr. Biol.* **2000**, *10*, 1247–1255. [[CrossRef](#)]
51. Liu, Y.; Wang, R.; Zhang, L.; Li, J.; Lou, K.; Shi, B. The lipid metabolism gene fto influences breast cancer cell energy metabolism via the pi3k/akt signaling pathway. *Oncol. Lett.* **2017**, *13*, 4685–4690. [[CrossRef](#)] [[PubMed](#)]
52. Redel, B.K.; Spate, L.D.; Lee, K.; Mao, J.; Whitworth, K.M.; Prather, R.S. Glycine supplementation in vitro enhances porcine preimplantation embryo cell number and decreases apoptosis but does not lead to live births. *Mol. Reprod. Dev.* **2016**, *83*, 246–258. [[CrossRef](#)] [[PubMed](#)]
53. Li, J.; Hou, W.; Zhao, Q.; Han, W.; Cui, H.; Xiao, S.; Zhu, L.; Qu, J.; Liu, X.; Cong, W.; et al. Lactate regulates major zygotic genome activation by h3k18 lactylation in mammals. *Natl. Sci. Rev.* **2024**, *11*, nwad295. [[CrossRef](#)] [[PubMed](#)]
54. Shi, W.; Cassmann, T.J.; Bhagwate, A.V.; Hitosugi, T.; Ip, W. Lactic acid induces transcriptional repression of macrophage inflammatory response via histone acetylation. *Cell Rep.* **2024**, *43*, 113746. [[CrossRef](#)]
55. Wang, H.; Cui, W.; Meng, C.; Zhang, J.; Li, Y.; Qian, Y.; Xing, G.; Zhao, D.; Cao, S. Mc1568 enhances histone acetylation during oocyte meiosis and improves development of somatic cell nuclear transfer embryos in pig. *Cell. Reprogramm.* **2018**, *20*, 55–65. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Genome-Wide Association Studies for Flesh Color and Intramuscular Fat in (Duroc × Landrace × Large White) Crossbred Commercial Pigs

Hao Li ^{1,2,†}, Cineng Xu ^{1,†}, Fanming Meng ^{2,†}, Zekai Yao ^{1,2}, Zhenfei Fan ¹, Yingshan Yang ¹, Xianglun Meng ¹, Yuexin Zhan ¹, Ying Sun ¹, Fucui Ma ¹, Jifei Yang ¹, Ming Yang ³, Jie Yang ^{1,4}, Zhenfang Wu ^{1,4,5}, Gengyuan Cai ^{1,4,*} and Enqin Zheng ^{1,*}

- ¹ College of Animal Science and National Engineering Research Center for Breeding Swine Industry, South China Agricultural University, Guangzhou 510642, China
 - ² State Key Laboratory of Livestock and Poultry Breeding, Guangdong Key Laboratory of Animal Breeding and Nutrition, Institute of Animal Science, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China
 - ³ College of Animal Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
 - ⁴ Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, South China Agricultural University, Guangzhou 510642, China
 - ⁵ Yunfu Subcenter of Guangdong Laboratory for Lingnan Modern Agriculture, Yunfu 527400, China
- * Correspondence: cgy0415@163.com (G.C.); eqzheng@scau.edu.cn (E.Z.)
- † These authors contributed equally to this work.



Citation: Li, H.; Xu, C.; Meng, F.; Yao, Z.; Fan, Z.; Yang, Y.; Meng, X.; Zhan, Y.; Sun, Y.; Ma, F.; et al. Genome-Wide Association Studies for Flesh Color and Intramuscular Fat in (Duroc × Landrace × Large White) Crossbred Commercial Pigs. *Genes* **2022**, *13*, 2131. <https://doi.org/10.3390/genes13112131>

Academic Editor: Bo Zuo

Received: 29 October 2022

Accepted: 12 November 2022

Published: 16 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The intuitive impression of pork is extremely important in terms of whether consumers are enthusiastic about purchasing it. Flesh color and intramuscular fat (IMF) are indispensable indicators in meat quality assessment. In this study, we determined the flesh color and intramuscular fat at 45 min and 12 h after slaughter (45 mFC, 45 mIMF, 12 hFC, and 12 hIMF) of 1518 commercial Duroc × Landrace × Large White (DLY) pigs. We performed a single nucleotide polymorphism (SNP) genome-wide association study (GWAS) analysis with 28,066 SNPs. This experiment found that the correlation between 45 mFC and 12 hFC was 0.343. The correlation between 45 mIMF and 12 hIMF was 0.238. The heritability of the traits 45 mFC, 12 hFC, 45 mIMF, and 12 hIMF was 0.112, 0.217, 0.139, and 0.178, respectively, and we identified seven SNPs for flesh color and three SNPs for IMF. Finally, several candidate genes regulating these four traits were identified. Three candidate genes related to flesh color were provided: *SNCAIP* and *PRR16* on SSC2, *ST3GAL4* on SSC5, and *GALR1* on SSC1. A total of three candidate genes related to intramuscular fat were found, including *ABLM3* on SSC2, *DPH5* on SSC4, and *DOCK10* on SSC15. Furthermore, GO and KEGG analysis revealed that these genes are involved in the regulation of apoptosis and are implicated in functions such as pigmentation and skeletal muscle metabolism. This study applied GWAS to analyze the scoring results of flesh color and IMF in different time periods, and it further revealed the genetic structure of flesh color and IMF traits, which may provide important genetic loci for the subsequent improvement of pig meat quality traits.

Keywords: meat quality; genome-wide association studies (GWAS); intramuscular fat; flesh color; pig; single nucleotide polymorphism (SNP)

1. Introduction

With the rise in consumption levels, consumers are increasingly demanding high pork quality, and they are willing to pay for high quality pork [1]. Meat quality is measured by flesh color, drip loss, shear force, cooking loss, intramuscular fat content (IMF), etc. [2,3]. The marbling and flesh color of pork usually depend on IMF [4], which refers to the fat entrapped in lean meat by muscle connective tissue, and it is usually regarded as the

quality standard of meats and is an important economic trait of domestic animals [5,6]. Therefore, the study of flesh color and IMF is of great significance and is influenced by genetics, nutrition, slaughtering method, and many other factors, of which genetic influence is great [7,8]. The study by Miar Y et al. [9] found that the heritability of flesh color and IMF were 0.20 and 0.23, respectively, and the heritability of flesh color and IMF was low to moderate. In recent years, many genes and single nucleotide polymorphism (SNP) loci related to meat quality have been detected using bioinformatics methods, of which genome-wide association studies (GWAS) have become a major method [10,11]. GWAS is an effective method used to analyze quantitative traits, and it has been widely used in the genetic analysis of quantitative traits in livestock and poultry [12,13]. Up until 24 August 2022, the pig quantitative trait locus (QTLs) database [14] included 70 QTLs for flesh color score (0.20% for), 890 QTLs (2.48%) for IMF, and 136 QTLs (0.38%) for marble score. The lack of these traits in pig QTL may be due to the complexity of the assay and the uncontrollable environment, which means that there are traits that still have the capacity for supplementation in this database. Salas et al. [15] found that pale soft exudative (PSE) meat was associated significantly with halothane genes, which is one of the causal genes for meat quality. Zhang et al. [16] demonstrated that the *PRKAG3* gene may have a significant impact on flesh color and pH. In addition, Wang et al. [17] found that *ZDHHC16*, *LOC102162218*, and *PCDH7* were significantly associated with IMF. Davoli et al. [18] identified *PPP3CA* as a potential candidate gene affecting IMF. Luo et al. [19] detected 36 SNPs on SSC12 using the 60k chip, and these SNPs are associated with IMF, marbling, and moisture [20–22]. Since previous studies have likely been mainly limited to population size or purebred status, resulting in large QTL spacing, and the causal genes have not been identified. In order to reveal the genetic mechanisms of meat quality, researchers have used the genome method to conduct in-depth analyses. Liu et al. [23] characterized the expression QTL (eQTL) in porcine skeletal muscle using RNA sequence data from 189 Duroc × Luchuan crossbred pigs, combined with genome-wide eQTL and allele specific expression (ASE) analysis that identified several new candidate genes for meat quality traits. In order to provide new genetic loci for the genetic improvement of meat quality traits and to better identify the QTL recombination of meat quality traits [24], we used a 1518 Duroc × Landrace × Large White (DLY) crossbred pig population to analyze the four traits (45 mFC, 12 hFC, 45 mIMF, and 12 hIMF) through GWAS, combined with the actual character correlation and heritability. We expect to provide new insight into the identification of candidate genes for flesh color and IMF and the molecular breeding of pigs.

2. Materials and Methods

2.1. Ethical Statement

All samples required for the experimental procedures were collected in this research according to the guidelines of Animal Care and Use by the China Agriculture and Rural Affairs Administration. The Animal Care and Use Committee of South China Agricultural University (SCAU) approved all animal experiments described in this study (2018F098).

2.2. Animals and Phenotypic Measurements

The experimental animals were provided by Guangdong Wens Foodstuff Group Co., Ltd. (Guangdong, China). A total of 1518 DLY pigs (757 boars and 764 sows) born from 2018 to 2019 were used for the experiment. These pigs were fattened from 3 farms and were divided into 13 batches for slaughter. The longissimus dorsi muscle between the 11th and 12th ribs was taken 45 min and 12 h after slaughter and scored for flesh color and IMF using the US NPPC (1991). The scoring was performed by three trained professionals and then averaged.

2.3. Genotyping and Quality Control

A total of 1518 ear tissues from the DLY pig was collected. Extraction of DNA from the ear tissues was completed using the traditional standard phenol/chloroform method, and the quality of the DNA samples was assessed using the light absorption ratio (A260/280)

of a spectrophotometer (Guangzhou Shenhua Biotechnology Co., Ltd., Guangzhou, China) and agarose gel electrophoresis [25,26]. Genotyping was performed according to the method described by Ding et al. [27]. According to the manufacturer's requirements, 1518 pigs were genotyped with the GeneSeek Porcine 50K SNP chip (Neogen, Lincoln, NE, USA) and quality filtering was completed using PLINK v1.90 [28]. Individuals with call rates of > 95% and markers with call rates of > 99%, a minor allele frequency (MAF) of > 99%, and a Hardy–Weinberg test p -value of > 10^{-6} were retained. In addition, all markers located on sex chromosomes and unmapped regions were excluded from analysis [29,30]. Briefly, a total of 28,066 SNPs were finally obtained for subsequent analysis after data editing. Furthermore, the heritability of the traits was estimated using GCTA v1.92 software in this study [31].

2.4. Single-Locus Genome-Wide Association Study

The GWAS refers to the method proposed by Zhou et al. [32] to correct the population structure, and an efficient mixed linear model (MLM) was used to analyze the effect of a single SNP on phenotypic traits to achieve a genetic algorithm [33–35]. GEMMA software was used to conduct correlation analysis on the meat quality traits in the study [36,37]. The four traits of meat quality were analyzed, respectively. The formula used in this experiment is as follows:

$$y = W\alpha + x\beta + \mu + \varepsilon;$$

$$u \sim MVN_n(0, \lambda\tau^{-1}K), \varepsilon \sim MVN_n(0, \tau^{-1}I_n)$$

where y is the phenotypic trait vector for all individuals, including 45 mFC, 12 hFC, 45 mIMF, and 12 hIMF. W is the correlation matrix based on the appropriate fixed effects, including the top five eigenvectors of the principal components analysis (PCA), slaughter batch, sex, farm, and liveweight; α is the vector of the corresponding coefficients, including the intercept; x is the vector of the SNP marker genotypes; β is the corresponding effective vector of the SNP markers; and μ and ε are the vectors of random effects and random errors, respectively. τ^{-1} is the variance of the random errors, λ is the ratio between the two variance components, K is the standard correlation matrix of the software estimation, I_n is an identity matrix, and MVN_n is a multivariate normal distribution. Since the Bonferroni correction is excessively stringent for false negative results, we refer instead to the false discovery rate (FDR) method [30,38,39] to limit the threshold of the GWAS results, and the formula for FDR screening is $P = FDR * X/N$, where the FDR is set to 0.01, X is the number of SNPs (with $p < 0.01$ in GWAS results), and N is the total number of SNPs that meet the requirements of the flesh color and intramuscular fat score. The thresholds of 45 mFC, 12 hFC, 45 mIMF, and 12 hIMF set by the FDR are 1.048×10^{-4} , 1.147×10^{-4} , 9.371×10^{-5} , and 9.691×10^{-5} , respectively.

2.5. Linkage Disequilibrium and Gene Annotation

QTL mapping was used for the linkage disequilibrium (LD) between the whole genome genetic markers and a specific quantitative trait designed to determine the locus affecting the trait and its position on a chromosome [40]. Haplotype block analysis was carried out using PLINK v1.90 [28] and Haploview [41] software. According Gebriel et al. [42], the search for genes was within the SNP locus and 500 Kb upstream and downstream of the GWAS-significant SNP [43] on Sscrofa11.1 in the Ensembl database (<http://asia.ensembl.org/index.html> accessed on 13 October 2022). Gene function annotation was performed on the screened genes, and KOBAS 3.0 [44] was used to further perform gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis on all of the screened genes.

3. Result and Discussion

3.1. Phenotypic Statistical Analysis

Table 1 shows the average scores of the DLY commercial pork color for 45 mFC, 12 hFC, 45 mIMF, and 12 hIMF. In our results, the phenotypes of the four traits measured

were all normally distributed (Figure S1). The population stratification was properly adjusted before the GWAS analysis and the genome inflation factor λ range was 0.989–1.047 (Figure 1). As shown in Table 1, the coefficients of variation (CV) for 45 mFC and 12 hFC were 27.6% and 42.2%, and the CVs of 45 mIMF and 12 hIMF were 41.1% and 42.1%, respectively. In addition, this study found that the heritability of both flesh color and IMF at 12 h was greater than that at 45 min, and the gap in the heritability of flesh color was more obvious. In comparison with previous studies [45], we obtained different results due to different measurement methods and breeds, suggesting that there is considerable potential for genetic improvement in flesh color and IMF. The genomic heritability of flesh color and IMF score was 0.112–0.217 (low to medium heritability). The flesh color heritability was lower than the 0.35 found by Zhang et al. [46] in the study of six populations (2448 pigs were tested for slaughter performance), and the difference in heritability may be due to the location of the phenotype measurement and the time of the postmortem evaluation. According to the American Pork Committee, the average subjective score of flesh color in the store was 2.85 ± 0.79 , the color score of the subjective in the laboratory was 2.74 ± 0.79 , the average subjective score of IMF in the store was 2.30 ± 1.07 , and the average subjective score of the IMF in the laboratory was 2.27 ± 1.02 [47], which is similar to our findings, where 45 mFC was larger than this data and the IMF was smaller than this data. As seen in Table 1, the average score of 45 mFC was 3.650 ± 1.006 , which is 1.5 higher than 12 hFC (2.193 ± 0.925). The decline of the score may have been caused by fat deposition, carotenoids, myoglobin content decreases, and the formation of lipid peroxides [48–50]. As long-term storage after slaughter causes a drop in pH, which created a visual difference for scorers, the oxidation of the freshly slaughtered pork formed an oxide layer that was visually pink and oxidized to brown over time, resulting in lower scores [47]. The IMF scores obtained in our experiment were similar in 45 m and 12 h, indicating that the subjective score of the IMF was less affected by the environment. The heritability of 45 mFC and 12 hFC were 0.112 and 0.139, respectively, and the phenotypic correlation between them was 0.343 ($p = 2.2 \times 10^{-16}$). The heritability of the IMF score at 45 m and 12 h was higher than that of flesh color, reaching 0.217 and 0.178, respectively, and the correlation between them was 0.238 ($p = 2.2 \times 10^{-16}$). It is worth noting that the deposition of the IMF affected the flesh color score; therefore, we estimated the correlation between flesh color and IMF at the same time, and the correlation between flesh color score and IMF score at 45 min was 0.135 ($p = 2.2 \times 10^{-16}$) and the correlation between them at 12 h was 0.293 ($p = 2.2 \times 10^{-16}$). This indicates that the two traits are somewhat correlated, with a slightly higher correlation at 12 h. Luo et al. [19] showed that the average scores of flesh color and IMF were 3.31 and 2.88, respectively, and the correlation between flesh color and IMF was 0.29, which was consistent with the correlation between the two traits at 12 h in our study. In conclusion, the flesh color scores in previous studies were consistent with those of our experiment, but the IMF scores were slightly higher than ours, which may be due to breed differences and personal score rating errors.

Table 1. Statistics of the phenotypic traits.

Phenotype ¹	Traits ²	Mean (SD) ³	CV ⁴	h^2 (se) ⁵	Phenotypic Correlation ⁶		p^7
Flesh color	45 mFC	3.650 ± 1.006	0.276	0.112 ± 0.032	0.343		2.2×10^{-16}
	12 hFC	2.193 ± 0.925	0.422	0.217 ± 0.042		0.135 (45 m)	1.4×10^{-7}
Intramuscular fat	45 mIMF	1.746 ± 0.717	0.411	0.139 ± 0.041	0.238	0.293 (12 h)	2.2×10^{-16}
	12 hIMF	1.515 ± 0.638	0.421	0.178 ± 0.041			2.2×10^{-16}

¹ Flesh color and IMF phenotypic traits; ² each trait is divided into different time points for statistics; ³ mean \pm standard deviation; ⁴ coefficient of variation; ⁵ heritability \pm standard error; ⁶ phenotypic correlations of the four traits (45 m is the correlation between flesh color and IMF at 45 min; 12 h is the correlation at 12 h); ⁷ phenotype-related p -values.

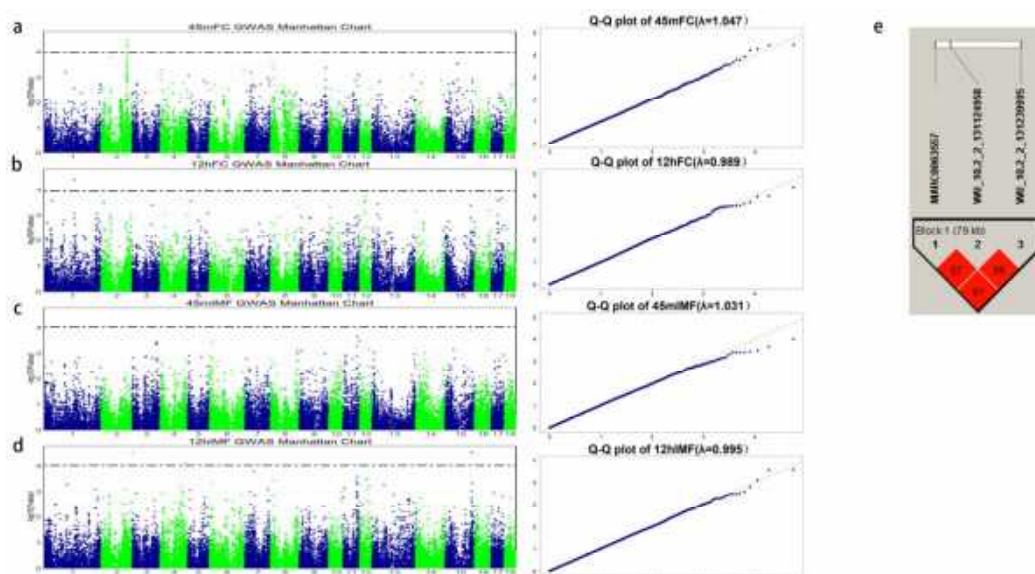


Figure 1. Manhattan plots of the single-locus GWAS and linkage disequilibrium. (a) 45-min GWAS analysis results of flesh color score. (b) IMF score 45-min GWAS analysis results. (c) 12-h GWAS analysis results of meat color score. (d) IMF score 12-h GWAS analysis results. (e) Haplotype blocks of top SNPs located on SSC2 for meat color traits. A haplotype block of 79 kb harbors the top SNP (shown in red and highlighted with black triangles). The x-axis represents chromosomes, the y-axis represents $-\log_{10}(p\text{-value})$, and the dashed line represents the FDR correction threshold. λ is the genome expansion coefficient, the abscissa is the quantile of the expected population distribution, and the ordinate is the quantile of the sample empirical distribution.

3.2. Flesh Color GWAS

Flesh color is the first perception of consumers, and it is largely determined by myoglobin, hemoglobin, and Fe^{2+} [49,51,52]. Because bloodletting after slaughter can lead to a loss of hemoglobin, we selected 45 min and 12 h after slaughter to evaluate the muscle color of the same pig, and then we recorded the phenotype for the GWAS. According to the method mentioned above, we used the FDR method for 45 mFC and 12 hFC for the statistics of the dominant loci. All the scores of 45 mFC and 12 hFC were statistically analyzed. Within the range set by the FDR, the number of SNPs in 45 mMC and 12 hMC was four and three, respectively, and the respective p -values were 1.048×10^{-4} and 1.147×10^{-4} . A total of seven SNPs was detected on the flesh color trait, of which four SNPs on 45 mMC were located in the 2 Mb (123.9–125.9 Mb) region on SSC2. We used haplotype software for the LD analysis of the SNPs, and there was a 79 Kb haplotype block near the top SNP (WU_10.2_2_13124958, Figure 1e). The gene closest to the top SNP (WU_10.2_2_131124958) was Synuclein α interacting protein (SNCAIP). DRGA0003514 on SSC2 was located at 201 Kb near Proline rich 16 (PRR16). PRR16 has been found to be able to control the size of mammalian cells, and the proline content was relatively high. The size of the cells was related to the level of intracellular proteins, and the proteins were related to flavor. Therefore, PRR16 can be considered as a gene related to meat quality [53]. There were three significant SNPs identified in 12 hFC (1.147×10^{-4}), of which the top SNP (ASGA0004802) was located on SSC1, with Galanin receptor 1 (GALR1) at 246 Kb away from it. The GAL family is a G protein-coupled receptor which affects the contraction and inflammation of myometrium in pigs, and it is also a glycine receptor. Glycine is one of the main flavor amino acids in meat [54]. The remaining SNP (WU_10.2_9_59299494) was located at SSC9, which was located at ST3 β -galactoside α -2,3-sialyltransferase 4 (ST3GAL4) within the interval. ST3GAL4 is involved in the synthesis of terminal N-glycans in various cells and tissues and is associated with sialic acid. In mammals, sialic acid is usually used as the terminal monosaccharide of glycans in glycoproteins and glycolipids [55]. The gene information and distance near these SNPs are listed in Table 2. As early as 1998, Andersson

et al. [56] used microsatellite markers to map QTLs for flesh color traits in European wild boars and large white pigs, and they mapped the QTLs affecting pork color on SSC2. In 2001, Malek et al. [57] located QTLs in 144.2–144.5 Mb on this chromosome, and a QTL interval affecting flesh color scores was determined by using the least square regression interval and molecular genome scanning. Subsequently, Rohrer et al. [58] performed genome scanning on the loci affecting meat quality traits in a Duroc × Landrace F2 population, and they found that the region of 151.0–155.0 Mb on SSC2 was related to flesh color scores, but we found that these studies did not overlap with our findings when we analyzed a large number of DLY pigs, and so we believe this is a new QTL interval that affects flesh color scores. The interval 123–127 Mb affecting the flesh color scores obtained can be used as a supplement to the QTL affecting pork meat quality traits. We found that the flesh color error rate measured at 45 min was smaller and the score was higher, and two genes (*SNCAIP* and *PRR16*) related to meat traits were found on SSC2.

Table 2. The four traits' SNP locations and gene interval information.

Phenotype ¹	SSC ₂	SNP	Position (bp) ³	P-Value ⁴	Gene	Range ⁵
45 mFC	2	WU_10.2_2_131124958	125,961,057	3.527×10^{-5}	<i>SNCAIP</i>	–4195
45 mFC	2	ALGA0015705	124,966,025	3.695×10^{-5}	<i>ENSSSCG00000041876</i>	within
45 mFC	2	DRGA0003514	123,998,374	5.468×10^{-5}	<i>PRR16</i>	+201,134
45 mFC	2	WU_10.2_2_130487175	125,006,692	6.348×10^{-5}	<i>ENSSSCG00000041876</i>	within
12 hFC	1	ASGA0004802	147,198,001	4.044×10^{-5}	<i>GALR1</i>	+245,820
12 hFC	5	ALGA0033448	88,285,420	1.042×10^{-4}	<i>ENSSSCG00000000905</i>	+7341
12 hFC	9	WU_10.2_9_59299494	53,460,098	1.106×10^{-4}	<i>ST3GAL4</i>	within
12 hIMF	2	H3GA0054148	150,317,009	2.728×10^{-5}	<i>ABLIM3</i>	within
12 hIMF	4	INRA0017040	117,256,570	8.211×10^{-5}	<i>DPH5</i>	+12,483
12 hIMF	15	WU_10.2_15_140224592	126,725,046	2.778×10^{-5}	<i>DOCK10</i>	–13,515

¹ 45 mFC, flesh color of 45 min; 12 hFC, flesh color of 12 h; 12 hIMF, intramuscular fat content of 12 h; ² SSC, sus scrofa chromosome; ³ SNP position in Ensembl; ⁴ p-value in GWAS; ⁵ +/– the SNP located upstream/downstream of the nearest gene.

3.3. IMF GWAS

The IMF score is based on the marble score, which shows a significant positive correlation [59] and is also tightly related to the breed [60]. The candidate genes for IMF include *ADRP* [61] and *CAST* [62]. Stewart et al. [63] used the relationship between marbling score and IMF to predict the palatability of beef to consumers, and IMF content and marbling were found to have remarkably similar precision. According to our previous calculation method, three SNPs in 12 hIMF were identified as significant loci (Table 2). However, no significant SNPs associated with IMF score measured at 45 min were found. It may be that some SNP sites had been deleted because the labeling density and conditions were too strict. Subsequently, new SNPs on IMF can be selected by increasing the labeling density or selecting a looser quality control condition [43]. Although only three SNPs were screened on 12 hIMF, both SSC2 (H3GA0054148) and SSC15 (WU_10.2_15_140224592) reached a loose genome-wide significant threshold (1/28066). An SNP on SSC2 was located on Actin binding LIM protein 3 (*ABLIM3*), which was identified as a candidate gene for IMF by GWAS in Chinese Lulai black pigs [64] associated with fat deposits. An SNP on SSC15 was found at 1.3 kb downstream of *DOCK10* in a GWAS of backfat thickness in Italian white pigs, and Fontanesi et al. [65] identified this gene as a candidate gene affecting subcutaneous fat at 126.5 Mb on SSC15, which is consistent with our findings. The *DPH5* gene was found on SSC4 and the *DPH5* protein is produced in *Escherichia coli*, and studies have shown that the elongation factor 2 of *DPH5* gene deletion mutants has weak ADP-ribose receptor activity [66] and may be related to the metabolism of the animal's body. In addition, the gene *GPC6* found by Ding et al. [27] using GWAS that affected IMF deposition was also validated, but it was not found in our study. We analyzed the interval (126–127 Mb) of the top SNP (WU_10.2_15_140224592) on SSC15 and found that the linkage degree between

WU_10.2_15_140224592 and MARC0011865 was relatively high ($r^2 = 0.50$). The top SNP showed moderate correlations with the remaining two SNPs, indicating that our research results on the IMF score have certain reference value.

3.4. Pathway Enrichment Analysis

We performed gene function enrichment analysis on the enriched genes using KOBAS 3.0 and enriched the four and six pathways on KEGG and GO, respectively (Table S1, Figure S2). On the KEGG enrichment pathway, *ADRB2* and *HTR4* were jointly enriched on the neuroactive ligand-receptor interaction, calcium signaling pathway, and cAMP signaling pathway, and two genes—*CSNK1A1* and *CUL3*—were enriched on the hedgehog signaling pathway. The Ca^{2+} signaling pathway is controlled by the increase and decrease in Ca^{2+} levels in the cytoplasm, and the regulation of Ca^{2+} levels in the cytoplasm depends on Ca^{2+} channels such as the endoplasmic reticulum, Golgi apparatus, and mitochondria [67]. Neuroactive ligand–receptor interaction can be related to neuronal activity, and it can enhance or inhibit neuronal activity, is related to cardiac function, regulates apoptosis, and may be related to activity ability and pigmentation [68]. The hedgehog signaling pathway is one of the key regulators of animal development, and the lack of this pathway can cause changes in mouse bones and muscles and affect metabolic capacity [69]. The exercise capacity of an animal affects the proportion of its muscle fibers, which, in turn, leads to changes in IMF, which affects its lean muscle percentage, resulting in changes in marbling [70]. In the GO functional enrichment analysis, these genes were enriched in pathways such as the positive regulation of protein ubiquitination, guanyl-nucleotide exchange factor activity, and protein homodimerization activity. The remaining pathways are mostly related to neural guidance and metabolism, which suggests that our intuitive impression of flesh color is not only guided by flesh color-related genes but is also the result of the coordinated expression of multiple genes.

4. Conclusions

In this study, the IMF and flesh color of 1518 DLY commercial pigs were analyzed in two time periods (45 min and 12 h) for GWAS, and we found SNPs significantly associated with flesh color scores in a haplotype block spanning 79 Kb on SSC2. The heritability of the traits 45 mFC, 12 hFC, 45 mIMF, and 12 hIMF were 0.112, 0.217, 0.139, and 0.178, respectively. Most of the genes related to flesh color scores were located in the 123–127 Mb range. Other genes related to IMF are located on SSC2, SSC4, and SSC15. A total of seven candidate genes (*SNCAIP*, *PRR16*, *ST3GAL4*, *GALR1*, *ABLM3*, *DPH5*, and *DOCK10*) were mapped for flesh color and IMF score, and we expect that seven genes related to meat quality traits can provide reference for the subsequent genetic improvement of meat quality.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/genes13112131/s1>, Figure S1: Four traits' phenotype normal distribution density. The horizontal axis is the phenotypic score and the vertical axis is the density. Table S1: GO and KEGG enrichment results. Figure S2: GO enrichment analysis result chart.

Author Contributions: G.C., Z.W. and E.Z. conceived and designed the experiment. H.L., C.X., Z.Y., Z.F., Y.Y., X.M. and Y.Z. collected the samples and performed the experiments. F.M. (Fucui Ma), J.Y. (Jifei Yang) and Y.S. recorded the phenotypes. F.M. (Fanming Meng), H.L. and E.Z. visualized the experimental results. J.Y. (Jie Yang) and M.Y. analyzed the data. H.L., C.X. and F.M. (Fanming Meng) wrote the manuscript. G.C. and E.Z. revised the manuscript for comments. Z.W. and J.Y. contributed to the materials. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Natural Science Foundation of Guangdong Province project (2018B030313011), the National Key Experiment of Livestock and Poultry Breeding Open Project Fund (2021GZ04), and the Guangdong Modern Agricultural Industrial Technology System pig innovation team project (2022KJ126).

Institutional Review Board Statement: All samples required for the experimental procedures were collected in this research according to the guidelines of Animal Care and Use by the China Agriculture

and Rural Affairs Administration. The Animal Care and Use Committee of South China Agricultural University (SCAU) approved all animal experiments described in this study (2018F098).

Informed Consent Statement: Not applicable.

Data Availability Statement: The data on the SNP chips of 1815 pigs are not released for the time being because they belong to commercial companies and are not for private use. If necessary, the corresponding author can be contacted to request them.

Acknowledgments: The authors would like to thank all staff at the pig core breeding farms of Wens Foodstuff Group Co., Ltd. (Guangdong, China), the State Key Laboratory of Pig Genetic Improvement and Production Technology, Jiangxi Agricultural University for their support in collecting the experimental samples, and all members of the National Engineering Research Center for Breeding Swine Industry for their help with the experimental data and methods.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Omana, D.; Goddard, E.; Plastow, G.; Janz, J.; Ma, L.; Anders, S.; Moore, S.; Bruce, H. Influence of on-farm production practices on sensory and technological quality characteristics of pork loin. *Meat Sci.* **2014**, *96*, 315–320. [[CrossRef](#)] [[PubMed](#)]
2. Watson, R.; Gee, A.; Polkinghorne, R.; Porter, M. Consumer assessment of eating quality-development of protocols for Meat Standards Australia (MSA) testing. *Aust. J. Exp. Agric.* **2008**, *48*, 1360–1367. [[CrossRef](#)]
3. Ji, J.; Zhou, L.; Huang, Y.; Zheng, M.; Liu, X.; Zhang, Y.; Huang, C.; Peng, S.; Zeng, Q.; Zhong, L.; et al. A whole-genome sequence based association study on pork eating quality traits and cooking loss in a specially designed heterogeneous F6 pig population. *Meat Sci.* **2018**, *146*, 160–167. [[CrossRef](#)] [[PubMed](#)]
4. Biswas, A.K.; Mandal, P.K. Chapter 1—Current perspectives of meat quality evaluation: Techniques, technologies, and challenges. In *Meat Quality Analysis*; Biswas, A.K., Mandal, P.K., Eds.; Academic Press: Cambridge, MA, USA, 2020; pp. 3–17. ISBN 978-0-12-819233-7.
5. Wood, J.; Nute, G.; Richardson, R.; Whittington, F.; Southwood, O.; Plastow, G.; Mansbridge, R.; da Costa, N.; Chang, K. Effects of breed, diet and muscle on fat deposition and eating quality in pigs. *Meat Sci.* **2004**, *67*, 651–667. [[CrossRef](#)] [[PubMed](#)]
6. Hocquette, J.F.; Gondret, F.; Baéza, E.; Médale, F.; Jurie, C.; Pethick, D.W. Intramuscular fat content in meat-producing animals: Development, genetic and nutritional control, and identification of putative markers. *Animal* **2010**, *4*, 303–319. [[CrossRef](#)]
7. Sellier, P. Genetics of meat and carcass traits. In *Genetics on Pig*; CABI: Wallingford, UK, 1998.
8. Franco, D.; Vazquez, J.A.; Lorenzo, J.M. Growth performance, carcass and meat quality of the Celta pig crossbred with Duroc and Landrace genotypes. *Meat Sci.* **2014**, *96*, 195–202. [[CrossRef](#)]
9. Miar, Y.; Plastow, G.S.; Moore, S.S.; Manafiazar, G.; Charagu, P.; Kemp, R.A.; Van Haandel, B.; Huisman, A.E.; Zhang, C.Y.; McKay, R.M.; et al. Genetic and phenotypic parameters for carcass and meat quality traits in commercial crossbred pigs¹. *J. Anim. Sci.* **2014**, *92*, 2869–2884. [[CrossRef](#)]
10. Schwab, C.R.; Baas, T.J.; Stalder, K.J.; Mabry, J.W. Effect of long-term selection for increased leanness on meat and eating quality traits in Duroc swine¹. *J. Anim. Sci.* **2006**, *84*, 1577–1583. [[CrossRef](#)]
11. Hu, Z.-L.; Park, C.A.; Reecy, J.M. Developmental progress and current status of the Animal QTLdb. *Nucleic Acids Res.* **2015**, *44*, D827–D833. [[CrossRef](#)]
12. Risch, N.; Merikangas, K. The Future of Genetic Studies of Complex Human Diseases. *Science* **1996**, *273*, 1516–1517. [[CrossRef](#)]
13. Qiao, R.; Gao, J.; Zhang, Z.; Li, L.; Xie, X.; Fan, Y.; Cui, L.; Ma, J.; Ai, H.; Ren, J.; et al. Genome-wide association analyses reveal significant loci and strong candidate genes for growth and fatness traits in two pig populations. *Genet. Sel. Evol.* **2015**, *47*, 17. [[CrossRef](#)] [[PubMed](#)]
14. Hu, Z.-L.; Park, C.A.; Reecy, J.M. Bringing the Animal QTLdb and CorrDB into the future: Meeting new challenges and providing updated services. *Nucleic Acids Res.* **2021**, *50*, D956–D961. [[CrossRef](#)] [[PubMed](#)]
15. Salas, R.C.D.; Mingala, C.N. Genetic Factors Affecting Pork Quality: Halothane and Rendement Napole Genes. *Anim. Biotechnol.* **2016**, *28*, 148–155. [[CrossRef](#)] [[PubMed](#)]
16. Zhang, C.; Wang, Z.; Bruce, H.; Kemp, R.A.; Charagu, P.; Miar, Y.; Yang, T.; Plastow, G. Genome-wide association studies (GWAS) identify a QTL close to PRKAG3 affecting meat pH and colour in crossbred commercial pigs. *BMC Genet.* **2015**, *16*, 33. [[CrossRef](#)] [[PubMed](#)]
17. Wang, X.W.; Ding, R.R.; Quan, J.P.; Yang, L.X.; Yang, M.; Zheng, E.Q.; Liu, D.W.; Cai, G.Y.; Wu, Z.F.; Yang, J. Genome-wide association analysis reveals genetic loci and candidate genes associated with intramuscular fat in Duroc pigs. *Front. Agric. Sci. Eng.* **2017**, *4*, 335–341. [[CrossRef](#)]
18. Davoli, R.; Luise, D.; Mingazzini, V.; Zambonelli, P.; Braglia, S.; Serra, A.; Russo, V. Genome-wide study on intramuscular fat in Italian Large White pig breed using the PorcineSNP60 BeadChip. *J. Anim. Breed. Genet.* **2015**, *133*, 277–282. [[CrossRef](#)] [[PubMed](#)]
19. Luo, W.; Cheng, D.; Chen, S.; Wang, L.; Li, Y.; Ma, X.; Song, X.; Liu, X.; Li, W.; Liang, J.; et al. Genome-Wide Association Analysis of Meat Quality Traits in a Porcine Large White × Minzhu Intercross Population. *Int. J. Biol. Sci.* **2012**, *8*, 580–595. [[CrossRef](#)]

20. Davoli, R.; Catillo, G.; Serra, A.; Zappaterra, M.; Zambonelli, P.; Zilio, D.M.; Steri, R.; Mele, M.; Buttazzoni, L.; Russo, V. Genetic parameters of backfat fatty acids and carcass traits in Large White pigs. *Animal* **2019**, *13*, 924–932. [\[CrossRef\]](#)
21. Liu, X.; Xiong, X.; Yang, J.; Zhou, L.; Yang, B.; Ai, H.; Ma, H.; Xie, X.; Huang, Y.; Fang, S.; et al. Genome-wide association analyses for meat quality traits in Chinese Erhualian pigs and a Western Duroc × (Landrace × Yorkshire) commercial population. *Genet. Sel. Evol.* **2015**, *47*, 44. [\[CrossRef\]](#)
22. Gao, G.; Gao, N.; Li, S.; Kuang, W.; Zhu, L.; Jiang, W.; Yu, W.; Guo, J.; Li, Z.; Yang, C.; et al. Genome-Wide Association Study of Meat Quality Traits in a Three-Way Crossbred Commercial Pig Population. *Front. Genet.* **2021**, *12*. [\[CrossRef\]](#)
23. Liu, Y.; Liu, X.; Zheng, Z.; Ma, T.; Liu, Y.; Long, H.; Cheng, H.; Fang, M.; Gong, J.; Li, X.; et al. Genome-wide analysis of expression QTL (eQTL) and allele-specific expression (ASE) in pig muscle identifies candidate genes for meat quality traits. *Genet. Sel. Evol.* **2020**, *52*, 59. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Jiao, R.; Chen, X.; Boerwinkle, E.; Xiong, M. Genome-Wide Causation Studies of Complex Diseases. *J. Comput. Biol.* **2022**, *29*, 908–931. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Stromberg, K.; Hurley, N.E.; Davis, N.L.; Rueckert, R.R.; Fleissner, E. Structural Studies of Avian Myeloblastosis Virus: Comparison of Polypeptides in Virion and Core Component by Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis. *J. Virol.* **1974**, *13*, 513–528. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Ozsensoy, Y.; Sahin, S. Comparison of different DNA isolation methods and use of dodecyl trimethyl ammonium bromide (DTAB) for the isolation of DNA from meat products. *J. Adv. Veter.- Anim. Res.* **2016**, *3*, 368. [\[CrossRef\]](#)
27. Ding, R.; Yang, M.; Quan, J.; Li, S.; Zhuang, Z.; Zhou, S.; Zheng, E.; Hong, L.; Li, Z.; Cai, G.; et al. Single-Locus and Multi-Locus Genome-Wide Association Studies for Intramuscular Fat in Duroc Pigs. *Front. Genet.* **2019**, *10*, 619. [\[CrossRef\]](#)
28. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [\[CrossRef\]](#)
29. Ding, R.; Yang, M.; Wang, X.; Quan, J.; Zhuang, Z.; Zhou, S.; Li, S.; Xu, Z.; Zheng, E.; Cai, G.; et al. Genetic Architecture of Feeding Behavior and Feed Efficiency in a Duroc Pig Population. *Front. Genet.* **2018**, *9*, 220. [\[CrossRef\]](#)
30. Zhuang, Z.; Ding, R.; Peng, L.; Wu, J.; Ye, Y.; Zhou, S.; Wang, X.; Quan, J.; Zheng, E.; Cai, G.; et al. Genome-wide association analyses identify known and novel loci for teat number in Duroc pigs using single-locus and multi-locus models. *BMC Genom.* **2020**, *21*, 344. [\[CrossRef\]](#)
31. Yang, J.; Manolio, T.A.; Pasquale, L.R.; Boerwinkle, E.; Caporaso, N.; Cunningham, J.M.; de Andrade, M.; Feenstra, B.; Feingold, E.; Hayes, M.G.; et al. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat. Genet.* **2011**, *43*, 519–525. [\[CrossRef\]](#)
32. Zhou, X.; Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **2012**, *44*, 821–824. [\[CrossRef\]](#)
33. Platt, R.W.; Leroux, B.G.; Breslow, N. Generalized linear mixed models for meta-analysis. *Stat. Med.* **1999**, *18*, 643–654. [\[CrossRef\]](#)
34. Yu, J.; Pressoir, G.; Briggs, W.H.; Vroh Bi, I.; Yamasaki, M.; Doebley, J.F.; McMullen, M.D.; Gaut, B.S.; Nielsen, D.M.; Holland, J.B.; et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **2006**, *38*, 203–208. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Hayes, B.J.; Goddard, M. Technical note: Prediction of breeding values using marker-derived relationship matrices. *J. Anim. Sci.* **2008**, *86*, 2089–2092. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Price, A.L.; Patterson, N.J.; Plenge, R.M.; Weinblatt, M.E.; Shadick, N.A.; Reich, D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **2006**, *38*, 904–909. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Zhou, X.; Stephens, M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* **2014**, *11*, 407–409. [\[CrossRef\]](#)
38. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300. [\[CrossRef\]](#)
39. Johnson, R.C.; Nelson, G.W.; Troyer, J.L.; Lautenberger, J.A.; Kessing, B.D.; Winkler, C.A.; O'Brien, S.J. Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genom.* **2010**, *11*, 724. [\[CrossRef\]](#)
40. Andersson, L. Genetic dissection of phenotypic diversity in farm animals. *Nat. Rev. Genet.* **2001**, *2*, 130–138. [\[CrossRef\]](#)
41. Barrett, J.C.; Fry, B.; Maller, J.; Daly, M.J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **2005**, *21*, 263–265. [\[CrossRef\]](#)
42. Gabriel, S.B.; Schaffner, S.F.; Nguyen, H.; Moore, J.M.; Roy, J.; Blumenstiel, B.; Higgins, J.; DeFelice, M.; Lochner, A.; Faggart, M.; et al. The Structure of Haplotype Blocks in the Human Genome. *Science* **2002**, *296*, 2225–2229. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Li, J.; Peng, S.; Zhong, L.; Zhou, L.; Yan, G.; Xiao, S.; Ma, J.; Huang, L. Identification and validation of a regulatory mutation upstream of the BMP2 gene associated with carcass length in pigs. *Genet. Sel. Evol.* **2021**, *53*, 94. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Bu, D.; Luo, H.; Huo, P.; Wang, Z.; Zhang, S.; He, Z.; Wu, Y.; Zhao, L.; Liu, J.; Guo, J.; et al. KOBAS-i: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res.* **2021**, *49*, W317–W325. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Khanal, P.; Maltecca, C.; Schwab, C.; Gray, K.; Tiezzi, F. Genetic parameters of meat quality, carcass composition, and growth traits in commercial swine. *J. Anim. Sci.* **2019**, *97*, 3669–3683. [\[CrossRef\]](#) [\[PubMed\]](#)

46. Zhang, Y.; Zhang, J.; Gong, H.; Cui, L.; Zhang, W.; Ma, J.; Chen, C.; Ai, H.; Xiao, S.; Huang, L.; et al. Genetic correlation of fatty acid composition with growth, carcass, fat deposition and meat quality traits based on GWAS data in six pig populations. *Meat Sci.* **2018**, *150*, 47–55. [\[CrossRef\]](#)
47. Larzul, C.; Lefaucheur, L.; Ecolan, P.; Gogué, J.; Talmant, A.; Sellier, P.; Le Roy, P.; Monin, G. Phenotypic and genetic parameters for longissimus muscle fiber characteristics in relation to growth, carcass, and meat quality traits in large white pigs. *J. Anim. Sci.* **1997**, *75*, 3126–3137. [\[CrossRef\]](#)
48. Buttle, L.; Crampton, V.; Williams, P. The effect of feed pigment type on flesh pigment deposition and colour in farmed Atlantic salmon, *Salmo salar* L. *Aquac. Res.* **2001**, *32*, 103–111. [\[CrossRef\]](#)
49. Wood, J.D.; Enser, M.; Fisher, A.V.; Nute, G.R.; Sheard, P.R.; Richardson, R.I.; Whittington, F.M.; Hughes, S.I. Fat deposition, fatty acid composition and meat quality: A review. *Meat Sci.* **2008**, *78*, 343–358. [\[CrossRef\]](#)
50. Lehnert, S.J.; Christensen, K.A.; Vandersteen, W.E.; Sakhrani, D.; Pitcher, T.E.; Heath, J.W.; Koop, B.F.; Heath, D.D.; Devlin, R.H. Carotenoid pigmentation in salmon: Variation in expression at *BCO2-1* locus controls a key fitness trait affecting red coloration. *Proc. R. Soc. B Biol. Sci.* **2019**, *286*, 20191588. [\[CrossRef\]](#)
51. Mancini, R.A.; Hunt, M.C. Current research in meat color. *Meat Sci.* **2005**, *71*, 100–121. [\[CrossRef\]](#)
52. Suman, S.P.; Joseph, P. Myoglobin Chemistry and Meat Color. *Annu. Rev. Food Sci. Technol.* **2013**, *4*, 79–99. [\[CrossRef\]](#)
53. Yamamoto, K.; Gandin, V.; Sasaki, M.; McCracken, S.; Li, W.; Silvester, J.L.; Elia, A.J.; Wang, F.; Wakutani, Y.; Alexandrova, R.; et al. Largen: A Molecular Regulator of Mammalian Cell Size Control. *Mol. Cell* **2014**, *53*, 904–915. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Jana, B.; Calka, J.; Miciński, B. Regulatory Influence of Galanin and GALR1/GALR2 Receptors on Inflamed Uterus Contractility in Pigs. *Int. J. Mol. Sci.* **2021**, *22*, 6415. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Feng, Q.; Tomoya, I.; Chengwei, D.; Jie, Y.; Yuqin, W.; Tomohiko, F.; Jianguo, G. ST3GAL3, ST3GAL4, and ST3GAL6 differ in their regulation of biological functions via the specificities for the α 2,3-sialylation of target proteins. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **2020**, *34*, 881–897.
56. Andersson-Eklund, L.; Marklund, L.; Lundström, K.; Haley, C.S.; Andersson, K.; Hansson, I.; Moller, M.; Andersson, L. Mapping quantitative trait loci for carcass and meat quality traits in a wild boar x Large White intercross. *J. Anim. Sci.* **1998**, *76*, 694–700. [\[CrossRef\]](#)
57. Malek, M.; Dekkers, J.C.; Lee, H.K.; Baas, T.J.; Prusa, K.; Huff-Lonergan, E.; Rothschild, M.F. A molecular genome scan analysis to identify chromosomal regions influencing economic traits in the pig. II. Meat and muscle composition. *Mamm. Genome* **2001**, *12*, 637–645. [\[CrossRef\]](#)
58. Rohrer, G.A.; Thallman, R.M.; Shackelford, S.; Wheeler, T.; Koohmaraie, M. A genome scan for loci affecting pork quality in a Duroc-Landrace F2 population. *Anim. Genet.* **2005**, *37*, 17–27. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Katsumata, M. Promotion of intramuscular fat accumulation in porcine muscle by nutritional regulation. *Anim. Sci. J.* **2011**, *82*, 17–25. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Plastow, G.; Carrión, D.; Gil, M.; García-Regueiro, J.; Furnols, M.F.; Gispert, M.; Oliver, M.; Velarde, A.; Guàrdia, M.; Hortós, M.; et al. Quality pork genes and meat production. *Meat Sci.* **2005**, *70*, 409–421. [\[CrossRef\]](#)
61. Chang, B.H.-J.; Li, L.; Saha, P.; Chan, L. Absence of adipose differentiation related protein upregulates hepatic VLDL secretion, relieves hepatosteatosis, and improves whole body insulin resistance in leptin-deficient mice. *J. Lipid Res.* **2010**, *51*, 2132–2142. [\[CrossRef\]](#)
62. Zhang, C.Y.; Wang, Z.; Bruce, H.L.; Janz, J.; Goddard, E.; Moore, S.; Plastow, G.S. Associations between single nucleotide polymorphisms in 33 candidate genes and meat quality traits in commercial pigs. *Anim. Genet.* **2014**, *45*, 508–516. [\[CrossRef\]](#)
63. Stewart, S.; Gardner, G.; McGilchrist, P.; Pethick, D.; Polkinghorne, R.; Thompson, J.; Tarr, G. Prediction of consumer palatability in beef using visual marbling scores and chemical intramuscular fat percentage. *Meat Sci.* **2021**, *181*, 108322. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Wang, Y.; Ning, C.; Wang, C.; Guo, J.; Wang, J.; Wu, Y. Genome-wide association study for intramuscular fat content in Chinese Lulai black pigs. *Asian-Australasian J. Anim. Sci.* **2019**, *32*, 607–613. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Fontanesi, L.; Schiavo, G.; Galimberti, G.; Calò, D.G.; Scotti, E.; Martelli, P.L.; Buttazzoni, L.; Casadio, R.; Russo, V. A genome wide association study for backfat thickness in Italian Large White pigs highlights new regions affecting fat deposition including neuronal genes. *BMC Genom.* **2012**, *13*, 583. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Mattheakis, L.C.; Shen, W.H.; Collier, R.J. DPH5, a methyltransferase gene required for diphthamide biosynthesis in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **1992**, *12*, 4026–4037. [\[PubMed\]](#)
67. Kong, F.; You, H.; Zheng, K.; Tang, R.; Zheng, C. The crosstalk between pattern-recognition receptor signaling and calcium signaling. *Int. J. Biol. Macromol.* **2021**, *192*, 745–756. [\[CrossRef\]](#)
68. Wang, J.; Cheng, J.; Zhang, C.; Li, X. Cardioprotection Effects of Sevoflurane by Regulating the Pathway of Neuroactive Ligand-Receptor Interaction in Patients Undergoing Coronary Artery Bypass Graft Surgery. *Comput. Math. Method. Med.* **2017**, *2017*, 3618213. [\[CrossRef\]](#)
69. Ingham, P.W.; Nakano, Y.; Seger, C. Mechanisms and functions of Hedgehog signalling across the metazoa. *Nat. Rev. Genet.* **2011**, *12*, 393–406. [\[CrossRef\]](#)
70. Song, S.; Cheng, H.; Jung, E.-Y.; Joo, S.-T.; Kim, G.-D. Muscle Fiber Characteristics on Chop Surface of Pork Loin (*M. longissimus thoracis et lumborum*) Associated with Muscle Fiber Pennation Angle and Their Relationships with Pork Loin Quality. *Korean J. Food Sci. Anim. Resour.* **2020**, *40*, 957–968. [\[CrossRef\]](#)

科学出版社
Science Press

ISSN 0253-9772
CODEN ICHUDW

遗传

Hereditas (Beijing)

YICHUAN 遗传

第4期
2020年 第42卷

Hereditas (Beijing)

● 中国精品科技期刊 ● 中文核心期刊 ● 中国科学引文数据库收录期刊 ● 美国MEDLINE收录期刊

第四十二卷

第四期

二〇二〇年四月



科学出版社

ISSN 0253-9772



中国遗传学会 主办
中国科学院遗传与发育生物学研究所

目次 | Contents

综 述

333 染色质转座酶可及性测序研究进展

吴杰, 全建平, 叶勇, 吴珍芳,

杨杰, 杨明, 郑恩琴

347 ATAC-seq 在复杂疾病研究中的应用进展

陈敏, 张峰, 孟紫媛, 张学军

354 NMD 逃逸机制及其在疾病治疗中的应用

程苗苗, 曹延廷

363 数字 PCR 技术及其在检测领域的应用

冯秀晶, 易红梅, 任星旭, 任佳丽,

葛建滔, 王风格

研究报告

374 中国人群中抗结核药物引发肝损伤的易感

基因标记研究

周晨希, 李沐, 怀聪, 贺林, 秦胜营

380 *tPA/gGH* 双基因转染山羊乳腺上皮细胞表达分析

宋绍征, 于康英, 张婷, 陆睿, 潘生强,

周鸣鸣, 成勇

388 乏情和发情初产母猪下丘脑-垂体-卵巢轴中

lincRNAs 表达谱比较分析

任巧玲, 张家庆, 陆东锋, 王璟, 陈俊峰,

马强, 白献晓, 郭红霞, 高彬文, 邢宝松

403 小拟南芥 *MKK* 基因家族全基因组鉴定及

进化和表达分析

李晓翠, 康凯程, 黄先忠, 范永斌, 宋苗苗,

黄韵杰, 丁佳佳

遗传

Hereditas (Beijing)

第 42 卷 第 4 期 2020 年 4 月



封面说明

自世界第一例克隆羊 Dolly 诞生以来, 利用体细胞克隆技术制备转基因动物乳腺生物反应器来生产医药蛋白, 一直是科学界研究的热点。人组织纤溶酶原激活剂(tissue-type plasminogen activator, tPA)是由血管内皮细胞合成并分泌的丝氨酸蛋白酶, 能够高效特异地溶解血栓, 是一种较好的溶栓药物。但是, 利用转基因克隆动物乳腺表达 tPA 的产量一直较低, 深入研究提高 tPA 的表达水平, 对大量生产高效溶栓药物至关重要。本期宋绍征等“*tPA/gGH* 双基因转染山羊乳腺上皮细胞表达分析”一文对山羊乳腺上皮细胞转染 *tPA/gGH* 双基因的诱导表达进行了系统的分析研究, 以期通过 *gGH* 基因协同作用促进 tPA 基因表达, 为将来制备高表达 tPA 转基因克隆山羊及其他乳腺生物反应器奠定基础。封面插图展示了以离体培养的乳腺上皮细胞作为供核细胞, 通过体细胞克隆技术制备的转基因克隆山羊, 证明了利用双转基因乳腺上皮细胞制备高表达克隆山羊的可能性。

染色质转座酶可及性测序研究进展

吴杰¹, 全建平¹, 叶勇¹, 吴珍芳¹, 杨杰¹, 杨明², 郑恩琴¹

1. 华南农业大学动物科学学院, 国家生猪种业工程技术研究中心, 广州 510642

2. 仲恺农业工程学院, 动物科技学院, 广州 510225

摘要: 染色质转座酶可及性测序(assay for transposase-accessible chromatin with high-throughput sequencing, ATAC-seq)诞生于 2013 年, 具有比脱氧核糖核酸酶 I 超敏感位点测序(deoxyribonuclease I hypersensitive site sequencing, DNase-seq)和微球菌核酸酶敏感位点测序(micrococcal nuclease sequencing, MNase-seq)更快速、灵敏、简便的优点, 是目前分析全基因组范围染色质开放区域的热点技术。通过该技术能获得染色质开放区域的相关信息, 从而映射出转录因子等调控蛋白的结合区域和核小体定位等信息, 对于研究表观遗传分子机制具有重要意义。本文比较了 5 种获取染色质开放区域技术的优缺点, 重点介绍了 ATAC-seq 的原理和主要流程, 描述了利用 ATAC-seq 技术研究染色质开放区域的发展概况以及 ATAC-seq 的相关应用, 期望对真核生物全基因组水平的染色质开放区域研究、顺式调控元件鉴定以及遗传调控网络的解析等提供借鉴。

关键词: 染色质转座酶可及性测序; 染色质开放区域; Tn5 转座酶; 表观遗传修饰; 转录因子

Advances in assay for transposase-accessible chromatin with high-throughput sequencing

Jie Wu¹, Jianping Quan¹, Yong Ye¹, ZhenFang Wu¹, Jie Yang¹, Ming Yang², Enqin Zheng¹

1. National Engineering Research Center for Breeding Swine Industry, College of Animal Science, South China Agricultural University, Guangzhou 510642, China

2. College of Animal Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China

Abstract: Assay for transposase accessible chromatin with high-throughput sequencing (ATAC-seq) was developed in 2013. It has the advantages of more convenient operation and higher efficiency for DNA recovery than DNase I hypersensitive site sequencing (DNase-seq) and micrococcal nuclease sequencing (MNase-seq). ATAC-seq currently is the most popular technique of genome-wide mapping for chromatin accessibility. It provides information on binding regions of

收稿日期: 2019-11-14; 修回日期: 2020-01-29

基金项目: 广东省“扬帆计划”引进创新创业团队项目(编号: 2016YT03H062), 广东省现代农业产业技术体系生猪创新团队项目(编号: 2019KJ126)和广东省自然科学基金项目(编号: 2017A030313213)资助[Supported by Guangdong YangFan Innovative and Entrepreneurial Research Team Program (No. 2016YT03H062), Guangdong Modern Agricultural Industry Technology System Pig Innovation Team Project (No. 2019KJ126) and Guangdong Natural Science Foundation (No. 2017A030313213)]

作者简介: 吴杰, 硕士研究生, 专业方向: 分子遗传与动物育种。E-mail: wujiezi163@163.com

通讯作者: 杨明, 博士, 高级畜牧师, 研究方向: 动物遗传育种。E-mail: yangming@zhku.edu.cn

郑恩琴, 硕士, 高级实验师, 研究方向: 遗传育种。E-mail: eqzheng@scau.edu.cn

DOI: 10.16288/j.ycz.19-279

网络出版时间: 2020/2/29 8:47:16

URI: <http://kns.cnki.net/kcms/detail/11.1913.R.20200228.0936.002.html>

transcription factors and nucleosome localization on the chromatin. Thus, ATAC-seq is of great significance for studying the epigenetics and molecular mechanisms in chromatin structure. In this review, we compare the advantages and disadvantages of multiple techniques for profiling chromatin accessibility, and summarize the principles, main process, development and applications of ATAC-seq. We hope this review will provide a reference for study of genome-wide mapping for chromatin accessibility, identification of cis-regulatory elements, and dissection of the epigenetic and genetic regulatory networks using the ATAC-seq technology in eukaryotes.

Keywords: assay for transposase-accessible chromatin with high-throughput sequencing; open chromatin regions; Tn5 transposase; epigenetic modification; transcription factor

自然界中的生物根据其细胞核类型可以分为原核生物和真核生物, 其中原核生物的细胞核无核膜包被, 其遗传物质 DNA 裸露在外; 而真核生物细胞的细胞核 DNA 并非裸露, 而是以左旋超螺旋的方式 (约 147 bp) 绕八聚体结构的组蛋白 1.67 圈, 进而形成核小体^[1,2]。相邻核小体的连接区由 10~80 bp 的游离 DNA 与组蛋白 H1 共同构成; 核小体通过连接区的连接形成串珠式结构, 这种串联结构进一步折叠、凝聚, 形成染色质; 最终多条染色质以高度螺旋化状态包裹于细胞核中^[3]。研究显示, 染色质开放区域的基因组占总 DNA 序列的 2%~3%, 且超过 90% 的开放区域均与转录因子(transcription factor, TF)的结合相关^[4]。以 TF 为代表的调控因子可与其他染色质结合蛋白相互作用, 从而动态调控和维持染色质稳态, 在发育过程的调控中发挥着不可替代的作用^[5~7]。在 DNA 复制或转录过程中, DNA 的折叠结构被打开, 一些染色质区域处于开放状态, 调控因子(如转录因子)会与这些裸露的无核小体结合的 DNA 部位结合, 进而调控 DNA 的复制或转录过程^[8]。此外, 有研究表明, DNA 折叠、凝聚形成的染色质物理结构并不是一成不变的, 仍然能够发生动态的表观遗传修饰, 如 DNA 甲基化、组蛋白修饰、染色质重塑等^[8~11]。因此, 通过了解相关获取染色质开放信息的技术, 学习技术原理和应用, 明确了这些技术对于基因组调控元件的鉴定、转录因子结合位点的识别及转录调控机制等研究均具有重要意义。本文主要综述了染色质可及性研究技术的发展概况、以及染色质转座酶可及性测序(assay for transposase accessible chromatin with high-throughput sequencing, ATAC-seq)技术的原理和应用, 以期表观遗传学研究提供重要的参考。

1 染色质开放区域研究技术的发展历程

染色质开放区域的研究源于人们发现某些染色质特定位点表现出对 DNase I 酶切的高度敏感性^[12~15]。后期研究表明, 这些 DNase I 敏感位点(deoxyribonuclease I hypersensitive site, DHS)通常是顺式调控元件所在区域^[16], 其染色质裸露、结构疏松, 可与转录因子结合, 从而便于 DNase I 与之结合并剪切, 进而表现出高度敏感性^[17]。基于上述原理, 染色质开放区域的鉴定工作也随之展开。最先开展的是 DHS 鉴定分析工作, 该分析依赖 DNase I 高度敏感性特点, 并与 Southern 杂交技术结合, 不过很快发现该方法的灵敏性和精确性都较低, 并且耗时费力^[18,19]。随着高通量测序技术(high-throughput sequencing, HTS)的发展及测序成本不断降低, 衍生出一系列研究染色质开放区域的技术与方法, 如脱氧核糖核酸酶 I 超敏感位点测序(deoxyribonuclease I hypersensitive site sequencing, DNase-seq)^[20]、微球菌核酸酶敏感位点测序(micrococcal nuclease sequencing, MNase-seq)^[11]、甲醛辅助性调控元件分离测序(formaldehyde-assisted isolation of regulatory elements followed by sequencing, FAIRE-seq)^[21]、核小体定位和甲基化组测序(nucleosome occupancy and methylome sequencing, NOMe-seq)^[22]和 ATAC-seq。在上述 5 种技术中, 获取染色质开放信息的方式分为 3 种: DNase-seq、MNase-seq 以及 ATAC-seq 采用酶切法; FAIRE-seq 采用物理断裂法; NOMe-seq 技术则利用甲基化修饰。5 种技术的具体信息见表 1。ATAC-seq 与其他 4 种技术相比表现出更为简便和高效的优势, 一经发明就被广泛采用, 成为当前染色

表 1 5 种染色质可及性研究技术介绍
Table 1 Introduction of five chromatin accessibility assays

技术	细胞类型及数量	获取方式	目的	特点	参考文献
DNase-seq	需要 1,000,000~10,000,000 的任何细胞类型	DNase I 酶切	获取染色质开放信息	(1)比传统方法操作更简便; (2)细胞需要量大; (3)酶最优酶切浓度确定过程繁琐; (4)样品制备过程复杂且耗时	[20,24]
MNase-seq	需要 1,000,000~10,000,000 的任何细胞类型	微球菌核酸酶酶切	绘制核小体图谱以间接探测染色质可及性	(1)操作简单, 后期数据处理方便; (2)细胞样本需要量大; (3)酶浓度和切割温度难以确定	[26,27]
FAIRE-seq	需要 100,000~10,000,000 的任何细胞类型	超声波物理断裂	获取染色质开放信息	(1)不用酶切、不需要分离出细胞核; (2)没有序列切割特异性; (3)细胞需要量大; (4)甲醛最佳交联程度难以确定	[21,24,31]
NOMe-seq	至少 1,000,000 的任何细胞类型	甲基化修饰	获得内源 DNA 甲基化的信息并定位核小体	(1)不需要使 DNA 断裂, 不会产生富集偏差; (2)同时获得含 GpC 和 CpG 二核苷酸的信息; (3)细胞需要量大	[22,34]
ATAC-seq	500~50,000 个新鲜分离的细胞	Tn5 转座酶酶切	获取染色质可及性、转录因子结合以及核小体定位信息	(1)过程简便、效率高; (2)数据分析工具不够成熟; (3)线粒体、叶绿体中的 DNA 污染; (4)冷冻组织细胞 DNA 提取效率低; (5) DNA 片段损失过多	[36~38]

质开放区域获取的前沿技术。下面将对上述染色质开放区域获取技术的发展历程、作用机理以及研究进程进行描述。

1.1 DNase-seq 及其衍生技术

早在 2006 年, DNase-seq 技术便被用于 DHS 区域的探究。该技术过程可简单概括为: 先通过裂解剂裂解细胞释放细胞核, 选用最佳浓度 DNase I 消化细胞核并包埋于低熔点凝胶琼脂糖塞中, 以减少额外的随机剪切; 随后, 将 DNA 片段平末端化并在两端连接上接头, 通过 PCR 扩增目的片段完成测序文库构建^[20]。DNase-seq 鉴定大多数转录因子的结合位点的精确性主要取决于酶切片段大小, 其中短片段(<100 bp)比较长片段效果更佳^[23]。DNase-seq 技术的建立使得 DHS 的分析鉴定变得更简便, 但存在细胞量需要量大(1,000,000~10,000,000)、样品制备过程复杂且耗时、最优酶浓度确定过程繁琐等缺点^[24]。2015 年, Cumbie 等^[25]发现使用传统 DNase-seq 技术消化拟南芥(*Arabidopsis thaliana*)时, 无法获得足够的 DNA 量用于建库测序。其主要原因在于植物细胞碎片及根毛等杂质占据了凝胶琼脂糖塞, 从而

导致 DNA 产量过低。因此, 他们在传统 DNase-seq 基础上建立了 DNase I-SIM (for simplified in-nucleus method)技术。该技术在 DNase I 消化细胞核前, 先采用 Percoll 层析液对细胞核进行初步纯化; 经 DNase I 消化后, 在 T4 DNA 聚合酶的作用下, DNA 片段的双链形成平末端, 而不包埋于琼脂糖塞中。这种改进的 DNase I-SIM 技术的优点在于, Percoll 梯度纯化后的细胞核及 T4 DNA 聚合酶处理后的 DNA 片段依然具备高度完整性, 在极大缩减 DNase-seq 文库制备时间的同时, 有效提升了 DNA 片段的回收效率。

1.2 MNase-seq 技术

MNase-seq 技术发明于 2008 年, 其原理与 DNase-seq 大体相似, 不同之处在于, 该技术采用的是微球菌核酸酶替代 DNase I 酶, 进而对细胞核中的 DNA 进行切割。该技术可通过揭示由核小体和其他调节因子占据的基因组区域, 从而间接探测染色质可及性并绘制核小体图谱^[26]。与 DNase-seq 相比, MNase-seq 具有操作简单、后期数据处理更方便等优点。然而, MNase-seq 技术仍存在一些技术弊端。

首先,其灵敏度同样受到细胞样本数量(1,000,000~10,000,000)、酶浓度和切割温度等因素的影响^[27];其次,微球菌核酸酶对 A/T 碱基序列存在切割倾向性^[28],从而无法精确切割核小体边界^[29]。

1.3 FAIRE-seq 及其衍生技术

2007 年, Giresi 等^[21]开发了一种物理打断 DNA 双链的技术,即 FAIRE-seq 技术,其过程相较于 DNase-seq 更为简便。该技术是利用超声波打断已用甲醛交联的 DNA 序列,在未解交联的条件下,通过酚-氯仿抽提,提取位于水相中的游离 DNA 并测序,最终获得相应的染色质开放区域信息^[21]。FAIRE-seq 具有不需要酶、不需要分离出细胞核、不受细胞类型限制、没有序列切割特异性^[30]、以及在增强子区域具有更高的覆盖率等优点^[31],但同样面临着样品需求量大的限制(需要 100,000~10,000,000 的细胞)^[24]。值得注意的是,该技术难以确定甲醛最佳交联程度,从而成为限制该技术应用的最大瓶颈。究其原因在于, DNA 与甲醛过度交联或不充分交联,均会影响到最终的测序结果^[32]。2009 年, Auerbach 等^[33]在此基础上创建了超声处理交联染色质测序(sonication of cross-linked chromatin sequencing, Sono-seq)技术,其原理与 FAIRE-seq 相同,其主要区别在于, FAIRE-seq 是在酚-氯仿抽提之后进行大小分级,选择特定大小范围如 100~350 bp 范围进行建库测序。由于存在大小分级选择这一关键步骤, Sono-seq 与 FAIRE-seq 各自所鉴定的 Peaks 存在明显的区别。

1.4 NOMe-seq 技术

NOMe-seq 技术由 Kelly 等^[22]于 2012 年发明,该技术利用 GpC 甲基转移酶(M.CviPI)通过甲基化修饰的方式处理开放区域的 GpC 二核苷酸。因为 GpC^m 不存在于人类基因组中, M.CviPI 使 GpC 甲基化为无内源背景的 GpC^m。随后,经过亚硫酸氢盐处理和全基因组测序,可同时获得含 GpC 和 CpG 二核苷酸的相关信息,从而能在全基因组范围内确定核小体的位置,同时还能获得内源 DNA 甲基化的信息^[22,34]。NOMe-seq 需要的细胞量为 1,000,000 个,材料处理因为不需要使 DNA 断裂,因此不会产生富集偏差,从而可降低假阳性的概率。但由于 NOMe-seq 不是基于先富集目的片段,然后再测序的方法,因

此需要大量的测序读长数据以获得足够的深度及基因组覆盖率,从而获取整个基因组的开放性水平^[35]。

1.5 ATAC-seq 技术

以上 4 种技术虽然都能应用于染色质开放性的表观基因组学研究,但通病是通常需要几万至数百万个细胞作为输入材料用以平均细胞群体的异质性,涉及到复杂、耗时的样品制备过程,且不能同时探究核小体定位、染色质可接近性和 TF 结合的相互作用。而多数情况下,很多重要且稀少的细胞亚型可能很难提供足够的样品量进行全基因组染色质可及性分析。2013 年, Buenrostro 等^[36]建立了材料需求量少、过程更为简便、效率更高的 ATAC-seq 技术。该技术仅需两步就能从 500~50,000 个细胞捕获染色质开放区域^[36,37]。与其他技术方法不同的是, ATAC-seq 利用高度活跃的 Tn5 转座酶代替 DNase I 核酸酶、微球菌核酸酶 MNase 等分析染色质可接近性,能够将目的 DNA 片段化、末端修复和加上测序所需的接头(adaptor)一步完成,从而使建库步骤变得极为简便,达到投入量更低、通量更高的建库效果。Tn5 转座子的深入研究以及高通量技术的快速发展使得 ATAC-seq 技术能够成功建立并广泛应用, ATAC-seq 技术以及其衍生技术如转座子超敏位点测序(transposome hypersensitive sites sequencing, THS-seq)、Omni-ATAC 技术使染色质开放区域的获取更加精准、高效,必将成为染色质开放区获取的主流技术之一。

2 ATAC-seq 作用机理和过程

2.1 Tn5 转座酶的发现及作用机理

20 世纪 40 年代,美国遗传学家 Barbara McClintock 发现了转座子(transposon)^[39]。转座子也叫跳跃基因或转座因子,是一段可以改变其在基因组中的位置的 DNA 序列。转座因子几乎存在于所有真核生物的基因组中,其衍生物构成了基因组的很大一部分,从而在基因组功能和进化过程中发挥着重要作用^[40]。根据转座子的结构特点和转座方式可将其分为 I 型和 II 型。I 型转座子,也称作 RNA 转座子(即反转录转座子),转座方式为“复制-粘贴”型,

即转座时先将自身转录获得 RNA,再反转录回 DNA,增加自身一倍的拷贝数,增加的 DNA 再转座至新的位置。II 型转座子,也称作 DNA 转座子,直接以“剪切-粘贴”的方式剪切下自身的 DNA 序列插入到新的位置,不会增加拷贝数。

Tn5 转座子是一种细菌转座子,属于 II 型转座子的一种^[41],最早是在 *Escherichia coli* 中被发现。Tn5 由编码卡那霉素(kanamycin, KAN)、新霉素(neomycin, NEO)、链霉素(streptomycin, STR)3 种抗生素的核心序列和位于侧翼的两个高度同源且倒置的 IS50(insertion sequence, IS)序列组成,该 DNA 序列全长 5818 bp^[42~45]。其中,IS50 序列可编码参与转座的蛋白:转座酶(transposase, Tnp)和转座阻遏蛋白(transposase inhibitor, Inh)。但由于左侧末端的 IS50L 序列的第 1442 位碱基处 T/A 碱基对被 C/G 碱基对取代发生突变,导致翻译提前终止,因此仅有 IS50R 能够表达正常有活性的 Tnp 和 Inh^[43]。每个 IS50 序列具有两个 19 bp 的倒置末端:外末端(outside end, OE)和内末端(inside end, IE)。两倒置末端有 7 个 bp 不同,该末端是 Tnp 的结合位点^[46~48]。

转座发生过程大致分为 3 个步骤:(1)形成转座复合体。Tnp 分子的两个 N 末端结构域分别结合到 Tn5 转座子的两个 OE 末端,形成两个 Tnp-OE 复合体^[49];随后以末端第 2 位序列为中心弯曲约 36°~48°^[50],两个复合体发生联会,Tnp 的 C 末端结构域相互作用而二聚体化,形成一个二聚体蛋白与两分子 DNA 组成的 Tn5 转座复合体^[51,52];(2) Tnp 切割。形成转座复合体结构后,Tnp 便具备了切割 DNA 的活性^[53],且正因为上述转座复合体结构,结合在左末端的 Tnp 便会催化右末端的磷酸二酯键水解,而结合在右末端的 Tnp 负责催化左末端的磷酸二酯键水解,以此有效防止 Tnp 只对转座子 DNA 链的一端进行切割^[54,55];(3)插入靶序列。Tnp 通过活化水分子水解 DNA 链,使 Tn5 两末端分别形成 3'-OH 亲核基团,该亲核基团对 DNA 互补链进行亲核攻击,形成发夹结构,另一活化的水分子水解发夹结构,使 Tn5 的两端均变为平末端,此时整个转座复合体离开供体 DNA,向靶序列结合^[56];转座子的 3'-OH 基团以交错的方式攻击靶序列中的磷酸二酯键,使转座子插入位点间形成 9 bp 的粘性末端,通过其 3'-OH 端同靶序列的 5'-P 之间形成共价键,插入到靶 DNA 中^[57,58],随后在 DNA 聚合酶的作用下,Tn5

的两侧翼形成 9 bp 的正向重复序列补齐缺口^[59]。至此,整个转座过程完成。

通过上述转座过程不难看出,体外 Tn5 转座过程仅需 4 个条件便能完成:Mg²⁺、转座子末端序列、Tnp 和靶 DNA^[60]。ATAC-seq 过程中使用的就是简化后的二聚体转座复合物。复合物仅含有 3 个部分:转座酶、末端序列和测序接头^[61],能够保证在切割 DNA 的同时连接上接头以便后续的测序工作。同时,简化复合物的 Tnp 在 Tn5 主链上携带了特异的点突变体,使 Tnp 具有了更高的活性^[60,62]。另外,之所以需要 Mg²⁺,是由于 Mg²⁺在转座过程中能协同亲核基团,在 Mg²⁺的作用下,转座酶上催化转座子运动的 DDE 基序(天冬氨酸和谷氨酸)与 Mg²⁺配位发生突变,使原本不活跃的转座酶变成高度活跃状态^[63,64],是完成转座必不可少的因子之一。目前,Tn5 转座子以其转座的随机性好、稳定性高、插入位点容易测序等特点,已经成为分子遗传学研究的热门工具^[65,66]。随着高通量测序技术的发展和实验通量的不断增加,Tn5 转座酶因其优势被应用得越来越广泛。其中,极速建库、长读长测序技术(single tube long fragment read, stLFR)、单细胞测序、Mate Pair 文库构建、染色质转座酶可及性可视化分析(assay of transposase-accessible chromatin with visualization, ATAC-seq),以及近几年发现的 Tn5 家族对于蛋白结合区域、互作基因片段等研究的帮助都显示 Tn5 转座酶拥有不可估量的应用潜力^[67~69]。Buenrostro 等^[36]建立的 ATAC-seq 技术,正是充分利用了 Tn5 酶在测序建库中的巨大优势,能高效、精准的从基因组水平鉴别出染色质开放区域,在生命科学领域的遗传学研究中发挥着至关重要的作用。

2.2 ATAC-seq 主要过程

ATAC-seq 涉及 3 个主要的步骤^[36,38]:(1)获取细胞核。使用冷裂解缓冲液裂解细胞;(2)转座和纯化(图 1A)。细胞核提取后立即将沉淀重悬于转座酶反应混合物中,转座后使用 Qiagen MinElute PCR Purification Kit 纯化样品;(3)PCR 扩增(图 1B)。纯化后,进行 qPCR 定量分析以及 PCR 扩增。上述过程大概需要 3 h,几处细节的处理尤为重要:(1)为减小 PCR 中的片段大小和 GC 偏差影响,需要通过 qPCR 来确定 PCR 后续的循环数,在饱和前停止扩增,以此保证转座后片段大小在 40 bp~1 kb 范围而

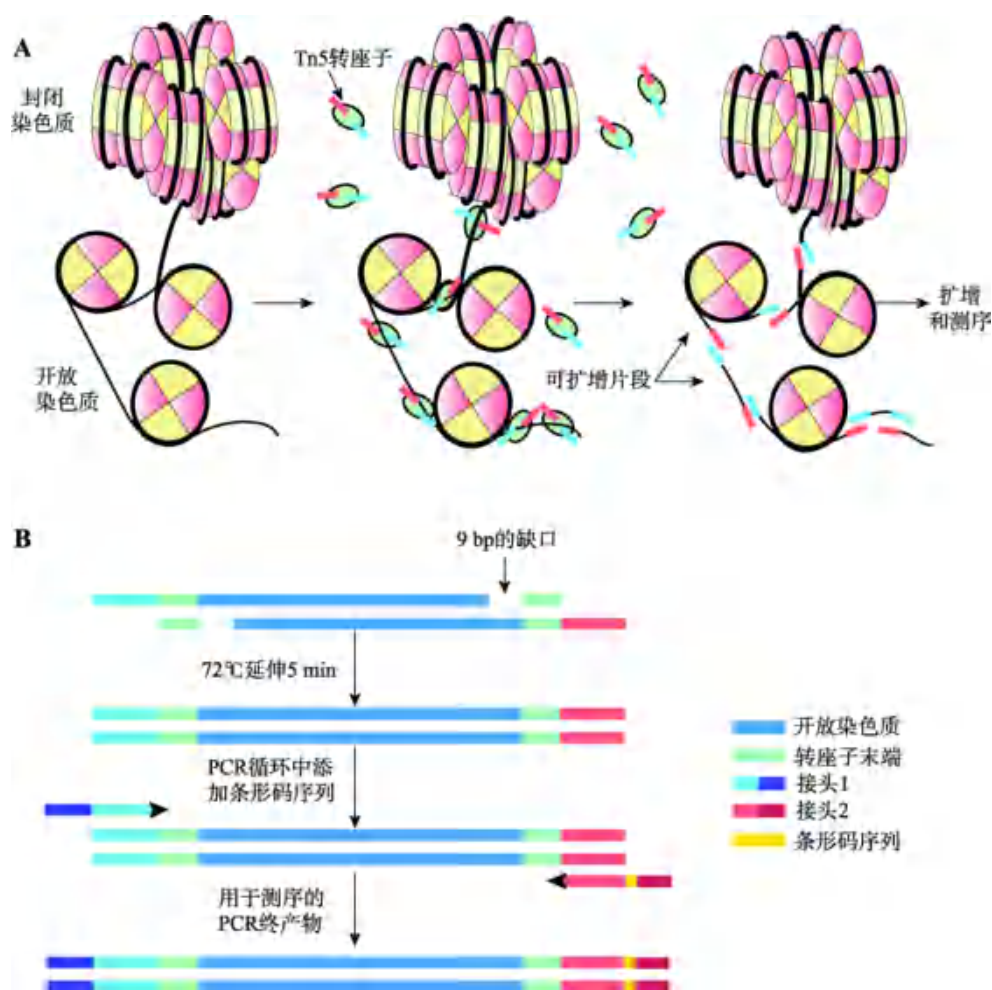


图 1 转座及扩增过程示意图

Fig. 1 The schematic transposition and PCR amplification process

A: 转座过程; B: 扩增过程。

不需要进行大小选择, 维持较高的库复杂性; (2) 因转座过程产生了 9 bp 的空隙, 因此 PCR 的第一步需要 72°C 反应 5 min 以填补该空隙, 且所用的 PCR 酶是具有链置换功能的非热启动酶; (3) 对于一个 DNA 片段而言, 两端接头连接是随机的。Tn5 酶切后会出现 3 类产物: 单端接头 1、单端接头 2 以及双端接头 1-接头 2, 只有连接不同接头的片段可用于富集扩增及测序。

2.3 数据分析

ATAC-seq 的数据分析过程大概可分为 4 个阶段^[24]: 第一阶段的数据预处理主要包括了过滤、比对和数据质量检测, 利用软件 CASAVA(Illumina) 获得 FASTQ 文件后过滤掉低质量片段, 由于相邻转座

的最小间隔为 38 bp, 通常 38 bp 以下的片段直接删除^[61], 并用 Bowtie 与参考基因组进行比对, 随后利用 SAMtools 去除重复的以及细胞器的 reads; 第二阶段是质控、数据可视化和 peak calling, 第二次的质控标准包括线粒体基因组所占比例高低和插入物大小分布图, 使用软件 IGV 可对数据进行可视化处理, 峰可以用软件 macs2、Hotspot 或者 ZINBA 寻找; 第三阶段是 peak 注释, 以获取基因组中 peak 的位置信息; 最后一步为模体(motif)注释和差异 peaks 分析, 将峰对应序列进行注释以间接确定转录因子信息, 同时, 利用 diffbind、DESeq2 等工具分析如不同的实验条件、多个时间节点、不同的发育时期等的差异区域, 最终获得转录因子结合位点的染色质可接近性状态信息。

2.4 ATAC-seq 优缺点分析

开展染色质开放区域的表现基因组学研究具有巨大的生物学意义,但过去的研究方法受到了复杂工作流程和大量细胞需求量的限制,从而导致该领域进展相对缓慢。直到 ATAC-seq 的出现,为注释开放染色质的基因组位置、DNA 结合蛋白、转录因子结合位点等基因组功能元件提供了新的契机。ATAC-seq 技术摆脱了像 DNase-seq 需要精确控制酶量以及 FAIRE-seq 需要确定甲醛交联时长等条件的限制,但依然存在影响其精确性的因素,如线粒体及植物细胞中叶绿体 DNA 的干扰、冷冻组织细胞 DNA 提取效率低、接头连接的随机性造成 DNA 片段的损失,以及大量酶切后的 DNA 片段过大而无法富集等^[24,36,38,70]。针对上述缺陷,同样产生了一系列改进措施。例如, Lu 等^[71]开发的 FANS-ATAC-seq (fluorescent activated nuclei sorting, FANS)、Roger 等^[72]开发的与细胞核基因组序列比对能达 90% 以上的 INTACT (isolation of nuclei tagged in specific cell types) 系统,以及与 INTACT 有相似结果的蔗糖沉淀法(crude)确保了在测定中使用高质量的完整细胞核的同时,能最大限度地减少线粒体和叶绿体中 DNA 的污染^[73]。Corces 等^[74]发明的 Omni-ATAC,提高了 ATAC-seq 对困难细胞系、稀少的原代细胞和临床上相关的冷冻组织中的应用普遍性。此外,针对接头的随机性和剪切后片段过大的问题, Sos 等^[75]开发了 THS-seq 技术,具有比传统 EzTn5 转座酶活性更高的新型 Tn5 超突变体(Tn5059)以及更优化的反应溶液和条件。同时设计 T7 启动子加转录引物替换原 Tn5 转座复合物中的 Adapter 1 和 2。通过转录生成单链 RNA,利用与 RNA 测序相同的原理获得 cDNA 并加上衔接子,最终完成建库。该技术避免了接头的随机连接,大大提高了转座效率,使得测序数据更为完整。随着 ATAC-seq 技术被不断改进, ATAC-seq 已逐渐成为目前染色质可及性分析的主流实验方法。

3 ATAC-seq 的应用和拓展

3.1 ATAC-seq 的应用

自 ATAC-seq 技术诞生起,该技术凭借其稳定性

和高灵敏度已广泛应用于表观基因组学研究。除了能用来确定功能基因组调控区域信息、找出组织特异基因以及预测潜在结合蛋白外,还能跟其他分析技术联合,如 RNA-seq、ChIP-seq(chromatin immunoprecipitation followed by high throughput sequencing)以及 Hi-C (high-through chromosome conformation capture)等,用以发现潜在的关键调控元件、转录因子和理解控制体内复杂过程的基因调控网络。其应用包括如下方面:

(1) DNA 调控功能元件注释。ATAC-seq 最直接的功能就是用来研究各种调控因子如 TF、启动子、增强子等的结合区域的开放状态以及对核小体进行定位,如对启动子区域开放状态的研究。Tan 等^[76]为探究寿命长且癌症发病率极低的裸鼹鼠的遗传机制,用 ATAC-seq 探测其染色质开放区域,结果显示裸鼹鼠的重编程基因的启动子区域更多的是处于关闭状态。通过对启动子区域开放状态后进行 SV40 LT (large t antigen, LT)抗原处理,发现 LT 可以抑制抑癌基因 *Rb*,从而提高重编程效率;而 LT 处理后,启动子区可接近性较处理前更高。上述结果表明裸鼹鼠细胞具有更稳定的表观基因组,提示可以利用这种稳定性为人类癌症预防和治疗提供新见解。同样,还有对转录因子结合区域开放性的研究。Scharer 等^[77]为深入了解系统性红斑狼疮(systemic lupus erythematosus, SLE)的表观遗传调控过程,对来自 SLE 和健康对照的 CD19⁺幼稚 B 细胞进行 ATAC-seq 分析。他们发现与健康对照相比,参与 B 细胞活化的基因周围的基因座,以及调节 B 细胞活化和分化的 TF 的结合位点在 SLE 细胞中可接近性更高,该结果验证了与 B 细胞活化相关的区域的开放状态,为进一步研究 SLE 疾病机制打下基础,充分体现了 ATAC-seq 确定染色质开放区域和鉴定顺式调控元件的能力。Kristofer 等^[78]利用 ATAC-seq 比较了致癌蛋白 Ras^{V12} 诱导的癌组织在早期肿瘤、晚期肿瘤与正常组织间染色质开放区域的状态。他们发现与正常组织比,肿瘤在发展过程中有数千个更容易接近的区域,从这些区域筛选并鉴定了在 Ras 依赖性肿瘤发生过程中异常活跃的调节区域,并结合 motif 分析确定出了结合这些区域的关键转录因子 AP-1 和 Stat92E。后期通过引入突变使 Stat92E 丧失功能后,发现肿瘤的严重程度降低,证明了转录因子

Stat92E 在肿瘤治疗发展中重要意义,也对 Ras 依赖性肿瘤的发展有了新的认知。上述基于 ATAC-seq 和 motif 分析的成功案例表明, ATAC-seq 可作为一种研究功能基因组调控区域和了解体内复杂基因调控网络的有效方式。另外,转录过程中,转录因子与核小体竞争结合 DNA 序列,转录因子结合处的核小体水平也因此较低^[79]。所以,获取核小体定位信息,对于了解转录调控、DNA 复制和修复等过程也很重要。Quillien 等^[80]为鉴定整个斑马鱼基因组中的细胞特异性增强子,对来自 Tg (fli1a: egfp)^{y1} 转基因胚胎的内皮细胞的细胞核进行 ATAC-seq 分析;在 FANS 技术的辅助下,用绿色荧光蛋白标记内皮细胞,通过荧光分离出细胞核获得高质量的 ATAC-seq 数据;后续通过分析短 DNA 片段(<100 bp)和长片段(180~247 bp)分别获得无核小体区及核小体结合区,以此定位核小体的位置。该研究揭示了整个基因组中转录起始位点的核小体定位模式以及与组蛋白修饰间的关联,还提供了在胚胎发育过程中控制基因表达的全基因组范围转录调控网络的动态信息。

(2) ATAC-seq 与 RNA-seq、ChIP-seq 等多组学数据的联合分析。ATAC-seq 技术更为巧妙的应用是将其获得的数据与其他表观遗传信息相结合,用以增强对科学问题的进一步解释。许多研究表明,通过联合 ATAC-seq 与 RNA-seq 数据进行分析,可发现潜在的特异性基因。Ackermann 等^[81]使用 ATAC-seq 对纯化的人 α 和 β 细胞中的开放染色质区进行了首次分析。通过与 RNA-seq 数据整合,进一步鉴定了两种细胞中已知的胰岛细胞转录因子的结合位点和已发现的 II 型糖尿病易感基因座的单核苷酸多态性(single nucleotide polymorphism, SNP)。更重要的是,发现了这两种细胞类型的新型特征基因,“组特异性蛋白质”(group specific protein)即维生素 D 结合蛋白仅存在于 α 细胞,而软骨素(chondrolectin)的免疫反应性仅存在于 β 细胞中。ATAC-seq 与 RNA-seq 的联合分析具备鉴定转录因子结合位点和发现潜在特异性基因的作用。2016 年, Garcia 等^[82]通过 ATAC-seq 比对分析了经典食道癌细胞系(OE33)、一种新发现的食管癌细胞系(MFD-1)、以及正常细胞系(HET1A)三种细胞系的特异染色质开放位点。他们发现和正常对照组相比, MFD-1 特异的染色质开放位点显著富集着 CTCF、NFY、Meis3、

Nrf2 的 motif,且这些位点附近基因富集在和食道癌及消化道肿瘤相关的功能基因区域内。后续结合全基因组测序和 RNA-seq 技术,发现了 MFD-1 的基因表达特性,进一步表明了 MFD-1 作为食管腺癌的临床疾病模型的可行性。同样,通过 ATAC-seq 与 RNA-seq 技术联合,可验证基因表达量与 ATAC-seq 信号的相关性,从而找出对应的转录因子。Ho 等^[83]为了找出特异表征间充质干细胞(mesenchymal stem cells, MSCs)的分子特征,从 8 周龄小鼠的四个组织(股骨和椎骨的骨髓、脂肪、肺)中分离出 MSCs 进行 ATAC-seq 和 RNA-seq 联合分析。首先,得出了 ATAC-seq 比转录组分析更加适合用来研究细胞特异性的结论,随后鉴定出可能有助于区分具有不同特征的 MSCs 的潜在转录因子,并推测这种通过研究染色质可及性来分析 MSCs 的方法,也可用于表征人类的 MSCs 和人类 MSCs 的临床应用。值得注意的是,虽然启动子可访问性和基因表达之间呈正相关,但也有许多研究表明,低表达或表达降低并不总是由于缺乏可访问区导致^[84,85]。因此, Starks 等^[86]再一次联合使用 ATAC-seq 和 RNA-seq 技术,探索染色质开放性与基因表达间的关联。试验通过分析妊娠中期小鼠胎盘的表达水平和启动子覆盖率将基因分成 3 组:高启动子覆盖率和表达的基因组(HA-HE)、中低覆盖率和表达的基因组(HA-ME)、低覆盖率和低表达的基因组(MA-ME)。结果发现 HA-HE 组基因可能是管家基因,而 HA-ME 组基因可能是组织特异性基因。随后,他们通过 motif 富集分析鉴定出了抑制 HA-ME 组基因的潜在转录因子,并发现 HA-ME 组的基因与胎盘的功能密切相关。该试验证明了 ATAC-seq 和 RNA-seq 的联合分析可用于小鼠和人的多种组织和细胞类型,用以鉴定活性抑制因子和组织特异性基因。

ATAC-seq 与 ChIP-seq 的联合使用常用于验证转录因子与目标开放区域的结合。2018 年, Rajbhandari 等^[87]联合 ATAC-seq、ChIP-seq 和 RNA-seq 揭示了抗炎因子 IL-10 的作用机理。为探究 IL-10 抑制脂肪细胞产热和能量输出的潜在机制,他们先利用 ATAC-seq 和 RNA-seq,验证了 IL-10 能改变脂肪细胞的染色质开放状态和降低产热相关基因 *UCP1* 的表达量。最后利用 ChIP-seq 揭示了 IL-10 通过抑制产热转录因子 ATF 和 C/EBP β 向增强子区

域募集进而达到抗炎作用的机制。Denny 等^[88]为探究小细胞肺癌的转移机制,联合 ATAC-seq、ChIP-seq 与 RNA-seq 对肺原发癌和肝脏转移癌进行比较分析。ATAC-seq 分析发现转移癌中的染色质开放性普遍增高,且差异的染色质开放性区域主要在基因远端调控元件上。随后通过 motif 富集分析、RNA-seq 和 ChIP-seq,发现显著富集的 *Nfib* 基因在转移癌中表达量增高,其 ChIP 信号与开放信号正相关,从而最终锁定关键基因 *Nfib*。

此外,还有联合 Hi-C 分析寻找调控元件如增强子。2018 年, Wang 等^[89]利用 ATAC-seq 和 Hi-C 研究多倍体棉花的三维基因组结构和转录调控之间的关系。他们发现一些由 DNase-seq 数据获得的 DHSs 与启动子之间存在互作,并推测这些处于开放状态的区域是潜在增强子,与 RNA-seq 数据整合进一步确认这些关键候选增强子具有转录活性。Mas 等^[90]联合 Hi-C、ATAC-seq、ChIP-seq 和 RNA-seq 四种技术,研究敲除甲基转移酶复合亚基基因 *MLL2* 后对基因组三维结构、染色质可接近性、组蛋白修饰以及表达水平的影响。首先利用 ChIP-seq 和 reChIP-seq 检测二价修饰,鉴定二价启动子。再通过 Hi-C 分析发现 *MLL2* 的敲除会引起二价基因转录起始位点与上下游互作模式的改变。随后 ATAC-seq 研究发现敲除 *MLL2* 引起二价启动子区的可接近性减少。最终揭示了 *MLL2* 通过作用于二价启动子,维持相对集中的二价基因间互作,可接近性状态和转录水平正常的作用。总之,ATAC-seq 能够与各种组学构建不同的关联模式,从不同的分析思路获取表观遗传信息或是探究转录调控机制,充分展现了其巨大的应用前景和无限的可能性。

3.2 单细胞 ATAC-seq 测序技术

染色质状态是以细胞类型特异性的方式动态调节的^[91]。虽然 DNase-seq、ATAC-seq 等技术用于测定全基因组水平染色质特定区域的可接近性,但测量获得的是群体细胞平均染色质状态,掩盖了细胞类型间和细胞内的异质性。因此,基于单细胞测序的 ATAC-seq (single-cell assay for transposase-accessible chromatin, scATAC-seq)被用于探测单细胞水平的染色质可及性^[92]。scATAC-seq 可应用于分析细胞亚群的基因调控网络、研究细胞异质性、发现生物

标志物、研究单细胞表观基因组学等^[93,94]。

目前 scATAC-seq 主要通过两种技术手段来高效获取单细胞全基因组范围染色质开放信息,分别为微流控技术^[92]和 ATAC-seq 组合标签方法 (single-cell combinatorial indexed ATAC-seq, sciATAC-seq)^[95]。单细胞 ATAC-seq 目前最常用的平台为 10× Genomics。其核心的微流控技术原理是将带有 barcode 信息的凝胶珠与转座酶处理后的细胞核混合,包裹在油滴中形成 GEMs (Gel Beads-in-emulsion)。一个特定 barcode 序列标记一个细胞核的所有序列以此区别各个细胞。单细胞 ATAC-seq 微流控技术因其捕获率高、容纳量大和价格相对较低等优点,被广泛应用于探索由表观遗传变化引起的细胞异质性、探索生物标志物、了解基因表达上游的基因调控网络等方面^[93]。另外,2015 年,华盛顿大学的 Jay Shendure 团队开发了 ATAC-seq 组合标签技术 sciATAC-seq^[95]。该方法不必依赖微流控平台,而是利用细胞标签技术对细胞核进行分子标记,通过两次稀释标记-混匀-再稀释标记-再混匀,使单个细胞能够被唯一标记而无需物理分离细胞。以此获得大量单细胞的染色质开放信息。该技术与微流控 scATAC-seq 相比,每个实验可共同测定数百万个单细胞,获得更多的单细胞信息。sciATAC-seq 的局限性是由于数据的稀疏性导致产生的数据集难以分析,无法获得较高的精确性^[96,97]。不过,相信随着数据集分析工具的不断改进,能够有效地提高分析精确性。如最新的分析工具 Scasat 可将开放的染色质信息作为二进制数据进行处理,并在保持数据的二进制特性的同时校正批次效应,使得该工具在分析 scATAC-seq 数据方面优于其他工具^[98]。sciATAC-seq 常用于获取单细胞的染色质调控信息以研究转录调控机制^[99,100]。

单细胞测序也同样发展到了单细胞水平的多组学整合分析,这对于准确解析细胞群中的细胞间差异至关重要。通过单细胞多组学联合分析,如通过联合 scATAC-seq 和 RNA-seq 技术,同时获得单细胞的表观基因组和转录组学信息,能够鉴定引起这些不同细胞表型的致病的顺式和反式作用元件,揭示基因表达调控特异性。如用于研究肿瘤异质性^[93]、揭示造血系统异质性^[101]等。Cusanovich 等^[94]为深入研究细胞亚群的基因表达调控机制,应用组合标签

技术 sciATAC-seq 分析 13 个成年小鼠组织的 10 万个单细胞的基因组范围内染色质的可接近性, 采用了一种 label-transfer 的方法, 整合 RNA-seq 和 ATAC-seq 数据对细胞进行聚类分析, 共鉴定出 30 个主要细胞亚群, 随后确定出 85 种不同的染色质可接近性模式, 以及近 40 万个差异可接近性元件。并使用这些数据将调节元件与其靶基因建立联系, 鉴定出了许多组织特异性的转录因子。之后, 通过将小鼠染色质开放性与人全基因组相关联, 揭示了部分人类遗传病与开放染色质间的潜在关系, 拓展了该研究的意义。整个实验为组织结构, 发育和分化, 各组织器官的调控网络的研究提供了丰富的参考。scATAC-seq 组合标签的方式进一步提高了单细胞 ATAC-seq 的通量, 标签组合作为单细胞基因组学的一种推广策略^[102], sciATAC-seq 具有不可估量的发展潜能。总之, 单细胞多组学联合分析无论是在疾病研究还是了解基因组功能等领域上都有广泛的应用前景。

4 结语与展望

鉴定染色质开放区域并对其进行精确定位, 对表观遗传学的研究具备重要意义。随着高通量测序技术的不断发展, 以 ATAC-seq 为代表的染色质开放区域获取技术, 将能系统发掘全基因组上的启动子、增强子、绝缘子和转录因子等重要调控元件的结合位点, 对深入了解整个基因调控网络具有重要意义。除了应用于寻找组织特异基因、定位核小体外, 将 ATAC-seq 的染色质开放信息进一步整合基因组、转录组、甲基化组等多组学数据, 可更加立体、直观地了解复杂基因之间的相互作用及其对表型的影响效应。

当前, ATAC-seq 技术已成为研究表观遗传调控的重要手段, 其更加简便的操作过程和更易满足的实验材料, 已在染色质开放区获取方面展现出无可比拟的优势和应用潜能。尽管该技术目前的数据分析工具还不够成熟, 但不容置疑的是, 其已成为表观遗传学研究的突破性技术。随着相应实验技术的进一步提高, 可以预期, ATAC-seq 将成为复杂性状遗传解析的研究利器之一, 从而进一步推动人类、小鼠及其他动植物等生命科学领域的稳步向前发展。

参考文献(References):

- [1] Kornberg RD. Chromatin structure: a repeating unit of histones and DNA. *Science*, 1974, 184(4139): 868–871. [\[DOI\]](#)
- [2] Richmond TJ, Finch JT, Rushton B, Rhodes D, Klug A. Structure of the nucleosome core particle at 7 Å resolution. *Nature*, 1984, 311(5986): 532–537. [\[DOI\]](#)
- [3] Zhou YB, Gerchman SE, Ramakrishnan V, Travers A, Muylldermans S. Position and orientation of the globular domain of linker histone H5 on the nucleosome. *Nature*, 1998, 395(6700): 402–405. [\[DOI\]](#)
- [4] Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, Garg K, John S, Sandstrom R, Bates D, Boatman L, Canfield TK, Diegel M, Dunn D, Ebersol AK, Frum T, Giste E, Johnson AK, Johnson EM, Kutayavin T, Lajoie B, Lee BK, Lee K, London D, Lotakis D, Neph S, Neri F, Nguyen ED, Qu H, Reynolds AP, Roach V, Safi A, Sanchez ME, Sanyal A, Shafer A, Simon JM, Song L, Vong S, Weaver M, Yan Y, Zhang Z, Zhang Z, Lenhard B, Tewari M, Dorschner MO, Hansen RS, Navas PA, Stamatoiyannopoulos G, Iyer VR, Lieb JD, Sunyaev SR, Akey JM, Sabo PJ, Kaul R, Furey TS, Dekker J, Crawford GE, Stamatoiyannopoulos JA. The accessible chromatin landscape of the human genome. *Nature*, 2012, 489(7414): 75–82. [\[DOI\]](#)
- [5] Poirier MG, Bussiek M, Langowski J, Widom J. Spontaneous access to DNA target sites in folded chromatin fibers. *J Mol Biol*, 2008, 379(4): 772–786. [\[DOI\]](#)
- [6] Fedor MJ. Chromatin structure and gene expression. *Curr Opin Cell Biol*, 1996, 4(18): 9384–9388. [\[DOI\]](#)
- [7] John S, Sabo PJ, Thurman RE, Sung MH, Biddie SC, Johnson TA, Hager GL, Stamatoiyannopoulos JA. Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nat Genet*, 2011, 43(3): 264–268. [\[DOI\]](#)
- [8] Bell O, Tiwari VK, Thomä NH, Schübeler D. Determinants and dynamics of genome accessibility. *Nat Rev Genet*, 2011, 12(8): 554–564. [\[DOI\]](#)
- [9] Kouzarides T. Chromatin modifications and their function. *Cell*, 2007, 128(4): 693–705. [\[DOI\]](#)
- [10] Jiang C, Pugh BF. Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet*, 2009, 10(3): 161–172. [\[DOI\]](#)
- [11] Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang ZB, Wei G, Zhao KJ. Dynamic regulation of nucleosome positioning in the human genome. *Cell*,

- 2008, 132(5): 887–898. [DOI]
- [12] Hewish DR, Burgoyne LA. Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochem Biophys Res Commun*, 1973, 52(2): 504–510. [DOI]
- [13] Scott WA, Wigmore DJ. Sites in simian virus 40 chromatin which are preferentially cleaved by endonucleases. *Cell*, 1978, 15(4): 1511–1518. [DOI]
- [14] Wu C, Bingham PM, Livak KJ, Holmgren R, Elgin SCR. The chromatin structure of specific genes: I. Evidence for higher order domains of defined DNA sequence. *Cell*, 1979, 16(4): 797–806. [DOI]
- [15] Stalder J, Larsen A, Engel JD, Dolan M, Groudine M, Weintraub H. Tissue-specific DNA cleavages in the globin chromatin domain introduced by DNase I. *Cell*, 1980, 20(2): 451–460. [DOI]
- [16] Mcghee JD, Wood WI, Dolan M, Engel JD, Felsenfeld G. A 200 base pair region at the 5' end of the chicken adult β -globin gene is accessible to nuclease digestion. *Cell*, 1981, 27(1, Part 2): 45–55. [DOI]
- [17] Gross DS, Garrard WT. Nuclease hypersensitive sites in chromatin. *Annu Rev Biochem*, 2003, 57(57): 159–197. [DOI]
- [18] Wu C. The 5' ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. *Nature*, 1980, 286(5776): 854–860. [DOI]
- [19] Crawford GE, Holt IE, Mullikin JC, Tai D, Blakesley R, Bouffard G, Young A, Masiello C, Green ED, Wolfsberg TG, Collins FS. Identifying gene regulatory elements by genome-wide recovery of DNase hypersensitive sites. *Proc Natl Acad Sci USA*, 2004, 101(4): 992–997. [DOI]
- [20] Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE. High-resolution mapping and characterization of open chromatin across the genome. *Cell*, 2008, 132(2): 311–322. [DOI]
- [21] Giresi PG, Kim J, Mcdaniell RM, Iyer VR, Lieb JD. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res*, 2007, 17(6): 877–885. [DOI]
- [22] Kelly TK, Liu YP, Lay FD, Liang GN, Berman BP, Jones PA. Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res*, 2012, 22(12): 2497–2506. [DOI]
- [23] He HH, Meyer CA, Hu SS, Chen MW, Zang C, Liu Y, Rao PK, Fei T, Xu H, Long H, Liu XS, Brown M. Refined DNase-seq protocol and data analysis reveals intrinsic bias in transcription factor footprint identification. *Nat Methods*, 2014, 11(1): 73–78. [DOI]
- [24] Tsompana M, Buck MJ. Chromatin accessibility: a window into the genome. *Epigenetics Chromatin*, 2014, 7(1): 33. [DOI]
- [25] Cumbie JS, Filichkin SA, Megraw M. Improved DNase-seq protocol facilitates high resolution mapping of DNase I hypersensitive sites in roots in *Arabidopsis thaliana*. *Plant Methods*, 2015, 11(1): 42. [DOI]
- [26] Rizzo JM, Sinha S. Analyzing the global chromatin structure of keratinocytes by MNase-seq. *Methods Mol Biol*, 2014, 1195: 49–59. [DOI]
- [27] Telford DJ, Stewart BW. Micrococcal nuclease: its specificity and use for chromatin analysis. *Int J Biochem*, 1989, 21(2): 127–137. [DOI]
- [28] Chung HR, Dunkel I, Heise F, Linke C, Krobisch S, Ehrenhofer-Murray AE, Sperling SR, Vingron M. The effect of micrococcal nuclease digestion on nucleosome positioning data. *PLoS One*, 2010, 5(12): e15754. [DOI]
- [29] Clark DJ. Nucleosome Positioning, Nucleosome spacing and the nucleosome code. *J Biomol Struct Dyn*, 2010, 27(6): 781–793. [DOI]
- [30] Zentner GE, Henikoff S. Surveying the epigenomic landscape, one base at a time. *Genome Biol*, 2012, 13(10): 250. [DOI]
- [31] Kumar V, Muratani M, Rayan NA, Kraus P, Lufkin T, Ng HH, Prabhakar S. Uniform, optimal signal processing of mapped deep-sequencing data. *Nat Biotechnol*, 2013, 31(7): 615–622. [DOI]
- [32] Simon JM, Giresi PG, Davis IJ, Lieb JD. Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. *Nat Protoc*, 2012, 7(2): 256–267. [DOI]
- [33] Auerbach RK, Euskirchen G, Rozowsky J, Lamarre-Vincent N, Moqtaderi Z, Lefrancois P, Struhl K, Gerstein M, Snyder M. Mapping accessible chromatin regions using Sono-Seq. *Proc Natl Acad Sci USA*, 2009, 106(35): 14926–14931. [DOI]
- [34] Rhie SK, Schreiner S, Farnham PJ. Defining regulatory elements in the human genome using nucleosome occupancy and methylome sequencing (NOME-Seq). *Methods Mol Biol*, 2018, 1766: 209–229. [DOI]
- [35] Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet*, 2019, 20(4): 207–220. [DOI]
- [36] Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Met*, 2013, 10(12): 1213–1218. [DOI]
- [37] Wu JY, Huang B, Chen H, Yin QZ, Liu Y, Xiang YL,

- Zhang BJ, Liu BF, Wang QJ, Xia WK, Li WZ, Li YY, Ma J, Peng X, Zheng H, Ming J, Zhang WH, Zhang J, Tian G, Xu F, Chang Z, Na J, Yang XR, Xie W. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature*, 2016, 534(7609): 652–657. [DOI]
- [38] Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol*, 2015, 109: 21–29. [DOI]
- [39] McClintock B. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci USA*, 1950, 36(6): 344–355. [DOI]
- [40] Bucher E, Reinders J, Mirouze M. Epigenetic control of transposon transcription and mobility in *Arabidopsis*. *Curr Opin Plant Biol*, 2012, 15(5): 503–510. [DOI]
- [41] Huang CR, Burns KH, Boeke JD. Active transposition in genomes. *Annu Rev Genet*, 2012, 46: 651–675. [DOI]
- [42] Berg DE, Davies J, Allet B, Rochaix JD. Transposition of R factor genes to bacteriophage lambda. *Proc Natl Acad Sci USA*, 1975, 72(9): 3628–3632. [DOI]
- [43] Rothstein SJ, Jorgensen RA, Postle K, Reznikoff WS. The inverted repeats of Tn5 are functionally different. *Cell*, 1980, 19(3): 795–805. [DOI]
- [44] Lovell S, Goryshin IY, Reznikoff WR, Rayment I. Two-metal active site binding of a Tn5 transposase synaptic complex. *Nat Struct Biol*, 2002, 9(4): 278–281. [DOI]
- [45] Auerswald EA, Ludwig G, Schaller H. Structural analysis of Tn5. *Cold Spring Harb Symp Quant Biol*, 1981, 45 Pt 1: 107. [DOI]
- [46] Berg DE. Julian Davies and the discovery of kanamycin resistance transposon Tn5. *J Antibiot*, 2017, 70(4): 339–346. [DOI]
- [47] Reznikoff WS. The TN5 transposon. *Annu Rev Microbiol*, 1993, 47(1): 945–963. [DOI]
- [48] Reznikoff WS, Bhasin A, Davies DR, Goryshin IY, Mahnke LA, Naumann T, Rayment I, Steiniger-White M, Twining SS. Tn5: A molecular window on transposition. *Biochem Biophys Res Commun*, 1999, 266(3): 729–734. [DOI]
- [49] Zhou M, Reznikoff WS. Tn5 transposase mutants that alter DNA binding specificity. *J Mol Biol*, 1997, 271(3): 362–373. [DOI]
- [50] York D, Reznikoff WS. DNA binding and phasing analyses of Tn5 transposase and a monomeric variant. *Nucleic Acids Res*, 1997, 25(11): 2153–2160. [DOI]
- [51] Steiniger-White M, Reznikoff WS. The C-terminal alpha helix of Tn5 transposase is required for synaptic complex formation. *J Biol Chem*, 2000, 275(30): 23127–23133. [DOI]
- [52] Bhasin A, Goryshin IY, Steiniger-White M, York D, Reznikoff WS. Characterization of a Tn5 pre-cleavage synaptic complex. *J Mol Biol*, 2000, 302(1): 49–63. [DOI]
- [53] Kale SB, Landree MA, Roth DB. Conditional RAG-1 mutants block the hairpin formation step of V(D)J recombination. *Mol Cell Biol*, 2001, 21(2): 459–466. [DOI]
- [54] Naumann TA, Reznikoff WS. Trans catalysis in Tn5 transposition. *Proc Natl Acad Sci USA*, 2000, 97(16): 8944–8949. [DOI]
- [55] Lisa AMB, Goryshin IY, Reznikoff WS. A mechanism for Tn5 inhibition. *J Biol Chem*, 1999, 274(1): 86. [DOI]
- [56] Bhasin A, Goryshin IY, Reznikoff WS. Hairpin formation in Tn5 transposition. *J Biol Chem*, 1999, 274(52): 37021–37029. [DOI]
- [57] Mizuuchi K, Adzuma K. Inversion of the phosphate chirality at the target site of Mu DNA strand transfer: Evidence for a one-step transesterification mechanism. *Cell*, 1991, 66(1): 129–140. [DOI]
- [58] Mizuuchi K. Transpositional recombination: mechanistic insights from studies of mu and other elements. *Annu Rev Biochem*, 1992, 61: 1011–1051. [DOI]
- [59] Crellin P, Chalmers R. Protein-DNA contacts and conformational changes in the Tn10 transpososome during assembly and activation for cleavage. *EMBO J*, 2001, 20(14): 3882–3891. [DOI]
- [60] Goryshin IY, Reznikoff WS. Tn5 *in vitro* transposition. *J Biol Chem*, 1998, 273(13): 7367–7374. [DOI]
- [61] Adey A, Morrison HG, Asan, Xun X, Kitzman JO, Turner EH, Stackhouse B, Mackenzie AP, Caruccio NC, Zhang X, Shendure J. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density *in vitro* transposition. *Genome Biol*, 2010, 11(12): R119. [DOI]
- [62] Reznikoff WS. Transposon Tn5. *Annu Rev Genet*, 2008, 42: 269–286. [DOI]
- [63] Davies DR, Mahnke Braam L, Reznikoff WS, Rayment I. The three-dimensional structure of a Tn5 transposase-related protein determined to 2.9-Å resolution. *J Biol Chem*, 1999, 274(17): 11904–11913. [DOI]
- [64] Peterson G, Reznikoff W. Tn5 transposase active site mutations suggest position of donor backbone DNA in synaptic complex. *J Biol Chem*, 2003, 278(3): 1904–1909. [DOI]
- [65] Reznikoff WS. Tn5 as a model for understanding DNA transposition. *Mol Microbiol*, 2003, 47(5): 1199–1206. [DOI]
- [66] Sakamoto H, Thiberge S, Akerman S, Janse CJ, Carvalho TG, Ménard R. Towards systematic identification of Plasmodium essential genes by transposon shuttle mutagenesis.

- nesis. *Nucleic Acids Res*, 2005, 33(20): e174. [DOI]
- [67] Caruccio N. Preparation of next-generation sequencing libraries using Nextera™ technology: simultaneous DNA fragmentation and adaptor tagging by in vitro transposition. *Methods Mol Biol*, 2011, 733: 241–255. [DOI]
- [68] Chen CY, Xing D, Tan LZ, Li H, Zhou GY, Huang L, Xie XS. Single-cell whole-genome analyses by linear amplification via transposon insertion (LIANTI). *Science*, 2017, 356(6334): 189–194. [DOI]
- [69] Chen XQ, Shen Y, Draper W, Buenrostro JD, Litzenburger U, Cho SW, Satpathy AT, Carter AC, Ghosh RP, East-Seletsky A, Doudna JA, Greenleaf WJ, Liphardt JT, Chang HY. ATAC-seq reveals the accessible genome by transposase-mediated imaging and sequencing. *Nat Methods*, 2016, 13(12): 1013–1020. [DOI]
- [70] Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf W. Transposition of native chromatin for multimodal regulatory analysis and personal epigenomics. *Nat Methods*, 2013, 10(12): 1213–1218. [DOI]
- [71] Lu Z, Hofmeister BT, Vollmers C, Dubois RM, Schmitz RJ. Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. *Nucleic Acids Res*, 2017, 45(6): e41. [DOI]
- [72] Deal RB, Henikoff S. A simple method for gene expression and chromatin profiling of individual cell types within a tissue. *Dev Cell*, 2010, 18(6): 1030–1040. [DOI]
- [73] Maher KA, Bajic M, Kajala K, Reynoso M, Pauluzzi G, West DA, Zumstein K, Woodhouse M, Bubb K, Dorrity MW, Queitsch C, Bailey-Serres J, Sinha N, Brady SM, Deal RB. Profiling of accessible chromatin regions across multiple plant species and cell types reveals common gene regulatory principles and new control modules. *Plant Cell*, 2018, 30(1): 15–36. [DOI]
- [74] Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, Kathiria A, Cho SW, Mumbach MR, Carter AC, Kasowski M, Orloff LA, Risca VI, Kundaje A, Khavari PA, Montine TJ, Greenleaf WJ, Chang HY. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods*, 2017, 14(10): 959–962. [DOI]
- [75] Sos BC, Fung HL, Gao DR, Osothpraprop TF, Kia A, He MM, Zhang K. Characterization of chromatin accessibility with a transposome hypersensitive sites sequencing (THS-seq) assay. *Genome Biol*, 2016, 17(1): 20. [DOI]
- [76] Tan L, Ke Z, Tomblin G, Macoretta N, Hayes K, Tian X, Lv R, Ablaeva J, Gilbert M, Bhanu NV, Yuan Z F, Garcia BA, Shi YG, Shi Y, Seluanov A, Gorbunova V. Naked mole rat cells have a stable epigenome that resists iPSC reprogramming. *Stem Cell Reports*, 2017, 9(5): 1721–1734. [DOI]
- [77] Scharer CD, Blalock EL, Barwick BG, Haines RR, Wei C, Sanz I, Boss JM. ATAC-seq on biobanked specimens defines a unique chromatin accessibility structure in naïve SLE B cells. *Sci Rep*, 2016, 6(1): 27030. [DOI]
- [78] Davie K, Jacobs J, Atkins M, Potier D, Christiaens V, Halder G, Aerts S. Discovery of transcription factors and regulatory regions driving in vivo tumor development by ATAC-seq and FAIRE-seq open chromatin profiling. *PLoS Genet*, 2015, 11(2): e1004994. [DOI]
- [79] Liu YJ, Zhang F, Liu HD, Sun X. The application of next-generation sequencing techniques in studying transcriptional regulation in embryonic stem cells. *Hereditas(Beijing)*, 2017, 39(8): 717–725.
刘亚军, 张峰, 刘宏德, 孙啸. 下一代测序技术在干细胞转录调控研究中的应用. *遗传*, 2017, 39(8): 717–725. [DOI]
- [80] Quillien A, Abdalla M, Yu J, Ou JH, Zhu LJ, Lawson ND. Robust identification of developmentally active endothelial enhancers in zebrafish using FANS-assisted ATAC-Seq. *Cell Rep*, 2017, 20(3): 709–720. [DOI]
- [81] Ackermann AM, Wang ZP, Schug J, Naji A, Kaestner KH. Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. *Mol Metab*, 2016, 5(3): 233–244. [DOI]
- [82] Garcia E, Hayden A, Birts C, Britton E, Cowie A, Pickard K, Mellone M, Choh C, Derouet M, Duriez P, Noble F, White MJ, Primrose JN, Strefford JC, Rose-Zerilli M, Thomas GJ, Ang Y, Sharrocks AD, Fitzgerald RC, Underwood TJ. Authentication and characterisation of a new oesophageal adenocarcinoma cell line: MFD-1. *Sci Rep*, 2016, 6: 32417. [DOI]
- [83] Ho YT, Shimbo T, Wijaya E, Ouchi Y, Takaki E, Yamamoto R, Kikuchi Y, Kaneda Y, Tamai K. Chromatin accessibility identifies diversity in mesenchymal stem cells from different tissue origins. *Sci Rep*, 2018, 8(1): 17765. [DOI]
- [84] Nelson AC, Mould AW, Bikoff EK, Robertson EJ. Mapping the chromatin landscape and Blimp1 transcriptional targets that regulate trophoblast differentiation. *Sci Rep*, 2017, 7(1): 6715–6793. [DOI]
- [85] Scott-Browne JP, López-Moyado IF, Trifari S, Wong V, Chavez L, Rao A, Pereira RM. Dynamic changes in chromatin accessibility occur in CD8⁺ T cells responding to viral infection. *Immunity*, 2016, 45(6): 1327–1340. [DOI]
- [86] Starks RR, Biswas A, Jain A, Tuteja G. Combined

- analysis of dissimilar promoter accessibility and gene expression profiles identifies tissue-specific genes and actively repressed networks. *Epigenetics Chromatin*, 2019, 12(1): 16. [DOI]
- [87] Rajbhandari P, Thomas B J, Feng A, Hong C, Wang JX, Vergnes L, Sallam T, Wang B, Sandhu J, Seldin M M, Lusis AJ, Fong LG, Katz M, Lee R, Young SG, Reue K, Smale ST, Tontonoz P. IL-10 signaling remodels adipose chromatin architecture to limit thermogenesis and energy expenditure. *Cell*, 2018, 172(1–2): 218–233.e17. [DOI]
- [88] Denny SK, Yang D, Chuang CH, Brady JJ, Lim JS, Grüner BM, Chiou S, Schep AN, Baral J, Hamard C, Antoine M, Wislez M, Kong CS, Connolly AJ, Park K, Sage J, Greenleaf WJ, Winslow MM. Nfib promotes metastasis through a widespread increase in chromatin accessibility. *Cell*, 2016, 166(2): 328–342. [DOI]
- [89] Wang MJ, Wang PC, Lin M, Ye ZX, Li GL, Tu LL, Shen C, Li JY, Yang QY, Zhang XL. Evolutionary dynamics of 3D genome architecture following polyploidization in cotton. *Nat Plants*, 2018, 4(2): 90–97. [DOI]
- [90] Mas G, Blanco E, Ballaré C, Sansó M, Spill YG, Hu D, Aoi Y, Le Dily F, Shilatifard A, Marti-Renom MA, Di Croce L. Promoter bivalency favors an open chromatin architecture in embryonic stem cells. *Nat Genet*, 2018, 50(10): 1452–1462. [DOI]
- [91] Stergachis AB, Neph S, Reynolds A, Humbert R, Miller B, Paige SL, Vernot B, Cheng JB, Thurman RE, Sandstrom R, Haugen E, Heimfeld S, Murry CE, Akey JM, Stamatoyannopoulos JA. Developmental fate and cellular maturity encoded in human regulatory DNA landscapes. *Cell*, 2013, 154(4): 888–903. [DOI]
- [92] Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, Chang HY, Greenleaf WJ. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*, 2015, 523(7561): 486–490. [DOI]
- [93] Liu LQ, Liu YC, Quintero A, Wu L, Yuan Y, Wang MY, Cheng MN, Leng LZ, Xu LQ, Dong GY, Li R, Liu Y, Wei XY, Xu JS, Chen XW, Lu HR, Chen DS, Wang QL, Zhou Q, Lin XX, Li GB, Liu SP, Wang Q, Wang HR, Fink JL, Gao ZL, Liu X, Hou Y, Zhu SD, Yang HM, Ye YM, Lin G, Chen F, Herrmann C, Eils R, Shang ZC, Xu X. Deconvolution of single-cell multi-omics layers reveals regulatory heterogeneity. *Nat Commun*, 2019, 10(1): 470. [DOI]
- [94] Cusanovich DA, Hill AJ, Aghamirzaie D, Daza RM, Pliner HA, Berletch JB, Filippova GN, Huang X, Christiansen L, Dewitt WS, Lee C, Regalado SG, Read DF, Steemers FJ, Disteche CM, Trapnell C, Shendure J. A single-cell atlas of *in vivo* mammalian chromatin accessibility. *Cell*, 2018, 174(5): 1309–1324. [DOI]
- [95] Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, Steemers FJ, Trapnell C, Shendure J. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science*, 2015, 348(6237): 910–914. [DOI]
- [96] Zamanighomi M, Lin ZL, Daley T, Chen X, Duren Z, Schep A, Greenleaf WJ, Wong WH. Unsupervised clustering and epigenetic classification of single cells. *Nat Commun*, 2018, 9(1): 2410. [DOI]
- [97] Schep AN, Wu BJ, Buenrostro JD, Greenleaf WJ. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods*, 2017, 14(10): 975–978. [DOI]
- [98] Baker SM, Rogerson C, Hayes A, Sharrocks AD, Rattray M. Classifying cells with Scasat, a single-cell ATAC-seq analysis tool. *Nucleic Acids Res*, 2019, 47(2): e10. [DOI]
- [99] Preissl S, Fang RX, Huang H, Zhao Y, Raviram R, Gorkin DU, Zhang YX, Sos BC, Afzal V, Dickel DE, Kuan S, Visel A, Pennacchio LA, Zhang K, Ren B. Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nat Neurosci*, 2018, 21(3): 432–439. [DOI]
- [100] Cusanovich DA, Reddington JP, Garfield DA, Daza RM, Aghamirzaie D, Marco-Ferreres R, Pliner HA, Christiansen L, Qiu X, Steemers FJ, Trapnell C, Shendure J, Furlong EEM. The cis-regulatory dynamics of embryonic development at single-cell resolution. *Nature*, 2018, 555(7697): 538–542. [DOI]
- [101] Buenrostro JD, Corces MR, Lareau CA, Wu B, Schep AN, Aryee MJ, Majeti R, Chang HY, Greenleaf WJ. Integrated single-cell analysis maps the continuous regulatory landscape of human hematopoietic differentiation. *Cell*, 2018, 173(6): 1535–1548. [DOI]
- [102] Ramani V, Deng X, Qiu R, Gunderson KL, Steemers FJ, Disteche CM, Noble WS, Duan Z, Shendure J. Massively multiplex single-cell Hi-C. *Nature Methods*, 2017, 14(3): 263–266. [DOI]



YICHUAN 遗传

1979 年创刊 月刊



主 编 张永清

主办单位 中国遗传学会; 中国科学院遗传与发育生物学研究所

刊载范围 遗传学、基因组学、发育生物学、生物进化及生物技术等领域有创新性的研究论文; 新技术与新方法; 学科热点问题的专论与综述; 遗传学教学的经验体会; 国内外著名遗传学家介绍; 国内外学术会议信息及科学新闻等

收录数据库 PubMed、MEDLINE、《生物学数据库》、《俄罗斯文摘杂志》、《化学文摘》、《中国学术期刊文摘》、《中国生物学文摘》、《中国生物医学文献数据库》等国内外相关文摘与检索系统收录

传播平台 中国遗传网、中国知网、万方数据、重庆维普、超星期刊域出版平台、生物通网站上全文发布

文章类型 特邀综述、综述、研究报告、技术与方法、资源与平台、遗传学教学、科学新闻、专家观点、争鸣与讨论、科学人生等

服务特色

- ★ 专业的学术评审。编委和评审专家均为国内遗传学领域的知名专家学者
- ★ 高效的审稿速度。网上投稿审稿, 免收审稿费, 评审时效 <4 周
- ★ 优先数字出版。录用文章即时在线发布, 网络出版时效 <100 天
- ★ 全文免费获取。中国遗传网提供文章免费下载服务
- ★ 资深的编辑服务。具十年以上工作经验的编辑提供专业出版服务; 英文特聘 UCSF 教授润色把关
- ★ 全面的推广传播。订户遍及国内高校、医院和科研院所
- ★ 精美的印刷装帧。出版后赠送作者精美样刊

投稿方式 请登录《遗传》网站 <http://www.chinagene.cn> 在“作者投稿查稿系统”注册投稿



欢迎投稿 欢迎订阅 欢迎刊登广告

刊号: ISSN 0253-9772, CN 11-1913/R 国内邮发代号: 2-810 国外发行代号: M62

定价: 80 元/期, 全年 960 元

地址: 北京市朝阳区北辰西路 1 号院 中国科学院遗传与发育生物学研究所 邮编 100101

电话: 010-64807669 E-mail: yczz@genetics.ac.cn

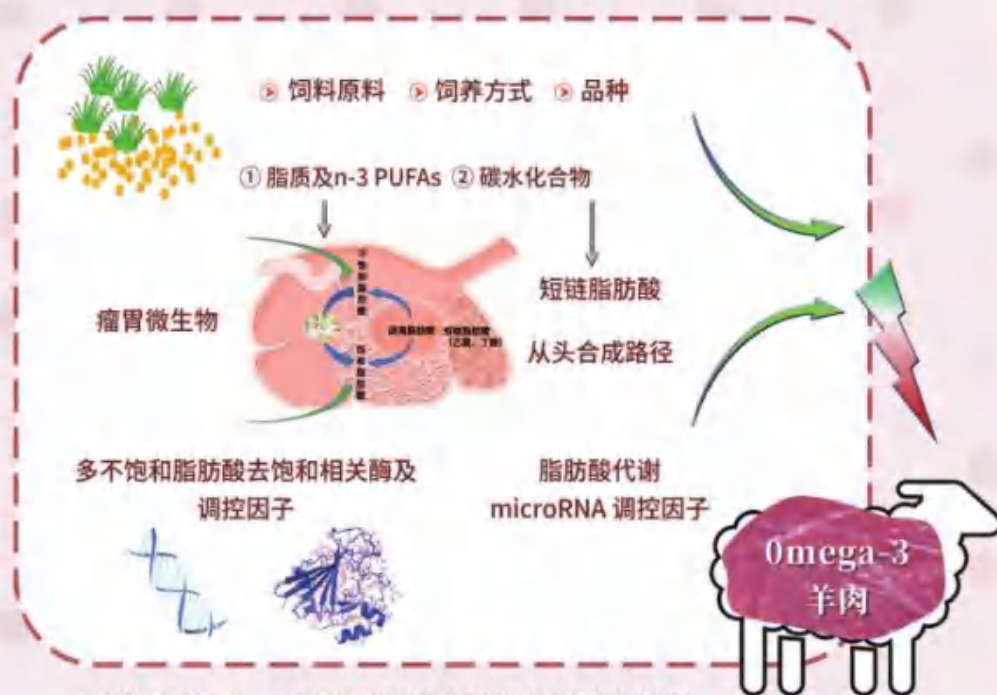
中国畜牧杂志

CHINESE JOURNAL OF ANIMAL SCIENCE

主管: 中国科学技术协会

主办: 中国畜牧兽医学学会

互动网站: www.bojar.cn

封面
文章n-3 多不饱和脂肪酸在
羊肉中沉积规律的研究进展

本文综述了羊肉中 n-3 多不饱和脂肪酸的来源及其在羊肉中的沉积规律,阐释了微生物及分子调控机制,为寻找提升羊肉中 n-3 多不饱和脂肪酸的沉积策略以及后续相关研究提供参考。



全国中文核心期刊
中国科技核心期刊
中国农林核心期刊

RCCSE中国权威学术期刊(A+)
中国学术期刊综合评价数据库来源期刊
中国学术期刊(网络版)来源期刊

2022. 10

10 / 2022

Started in 1953

Supervised by China Association for Science and Technology

Sponsored by China Association for Animal Science and
Veterinary Medicine

Published by Chinese Journal of Animal Science and
Veterinary Medicine Co. Ltd

Editor-in-Chief: Guo Yuming

Address: Beijing Boyar Communication Co, Ltd

Room C01, 02, #6 Building, Institute of Animal Science of
CAAS, No. 2 Yuanmingyuan West Road, Haidian District,
Beijing, 100193, P. R. China

Tel: (010)62732723 62734608
(010) 82893431

Fax: (010)82894198

E-mail: zgxmzz@cau.edu.cn (Contribution)
editor@boyar.cn (Advertisement)

Website: <http://www.zgxmzz.cn>

Distributor: China International Book Trading Corp

Oversea P. O. Registration No. SM452

Subscription Rate: US \$ 40.00 (per copy)

中国畜牧杂志

CHINESE JOURNAL OF ANIMAL SCIENCE

1953 年创刊

主 管 中国科学技术协会
主 办 中国畜牧兽医学会
学术支撑单位 中国农业大学动物科学技术学院
出 版 单 位 《中国畜牧兽医杂志》有限公司
地 址 北京市海淀区圆明园西路2号院56号
楼1层101、102
邮 编 100193
电 话 (010) 62732723 62734608
(010) 82893431
传 真 (010) 82894198
期 刊 网 址 <http://www.zgxmzz.cn>
(投稿系统)
互 动 网 站 <http://www.boyar.cn>
电 子 邮 箱 zgxmzz@cau.edu.cn (稿件咨询)
editor@boyar.cn (广告咨询)
广告总代理 北京博亚和讯文化传媒有限公司
电 话 (010) 82894056 82894057
传 真 (010) 82894198
发 行 电 话 (010) 82893871
出 版 日 期 2022 年 10 月 10 日
订 阅 全国各地邮局(所)
国内总发行 北京报刊发行局
国内邮发代号 82-147
国外总发行 中国国际图书贸易总公司
国外发行代号 SM452
中国标准刊号 ISSN 0258-7033
CN 11-2083/S
广告发布登记 京海工商广登字 20170255 号
印 刷 北京科信印刷有限公司
国 内 定 价 50 元

《中国畜牧杂志》编委会

(以姓名笔画为序)

主任

吴常信

副主任

冯子明 杨汉春 谢波平

顾问

印遇龙 刘守仁 李德发 张 沅
陈建春 尚志标 姚 斌
黄路生

编委会成员

刁其玉 马月辉 王立贤 王桂华
王洪荣 王泽茂 王 皓 王楚琳
史 杰 尹瑞春 郭平涛 尔声耀
冯定远 曲鲁江 朱化娜 乔玉峰
刘玉满 刘庆平 刘仲华 刘国世
刘建朝 刘剑锋 刘继军 齐广海
齐德生 何之春 孙永旋 李学伟
李 奎 李保明 李胜利 李 辉
杨小军 杨 宁 杨红建 连正兴
吴中红 张万刚 张守全 张发楠
张新军 张兴杰 张淑敏 张爱忠
张 浩 张旭杰 陈代文 陈国宏
陈显生 郭彩梅 林 亦 罗海玲
金立志 赵广永 赵书江 赵国琦
侯水生 侯永清 侯军威 姚军虎
贺 鑫 秦玉昌 秦应和 秦贵仁
袁长全 袁志刚 高玉娟 曹兵海
曹建民 康相泽 葛 翔 蒋宗勇
韩彦明 陆明星 曾申明 曾建国
谭全明 蔡祥盛 高仕彦

主编

冯子明

常务副主编

杨汉春

副主编

尹靖东 曲鲁江 刘国世 李保明
杨红建 张万刚 张爱忠 张 浩
贺 鑫 曹建民 霍立军

编辑部主任

孔平涛

编辑部副主任

赵 楠

市场总监

朱晓娟

总策划

王必勇

责任编辑

李蕊兰 赵 楠 郑永艳 周会会

综 述

瘦肉型猪基因组中 ROH 的应用及研究进展

李桂新, 王诗福, 杨 杰, 王小鹏, 郑恩恩 (13)

不饱和脂肪酸瘤胃微生物氢化与调控奶牛泌乳性能的研究进展

曹宜宜, 李秋爽, 王 敏, 谭友良 (8)

姜黄素的抗氧化功能及其在畜禽养殖中的应用研究进展

钟江文, 张海波, 欧阳红兵, 蔡力之, 毕晓娟, 吴伟斌, 郭冬生 (14)

n-3 多不饱和脂肪酸在羊肉中沉积规律的研究进展

张利阳, 张得康, 罗海玲, 王 娟 (19)

低蛋白氨基酸平衡日粮的应用及其对猪肉品质影响的研究进展

杨景森, 王 娟, 游宏勇, 朱 琴 (24)

奶牛不同体液中表征氧化应激生物标志物的研究进展

郭 琳, 葛 杰, 邓金莹, 刘建新, 王德伟 (31)

饲料添加剂包被产品的制备和检测方法研究进展

吴 鹏, 喻二华, 赵 彪, 侯 磊, 张倩云, 王 敏, 李 苗, 倪海洋, 程翠琴, 李海峰 (37)

单细胞测序技术及其在畜禽遗传育种中的应用研究新进展

赵 彪, 曹可欣, 张子敬, 刘 慧, 蔡翠翠, 李欣梅, 雷初朝, 陈 宏, 曹永霖 (44)

MSRN 基因突变对肌肉牛难产现象分析

杨 磊, 郝振廷, 白春玲, 李光昭 (49)

m⁶A 甲基化修饰及其影响动物脂肪生成的分子机制研究进展

宋开强, 赵 磊, 刘 慧, 姚 治, 张子敬, 雷初朝, 王二耀, 曹永霖 (53)

姜黄素的提取、生理功能及其在肉鸡生产中的应用研究进展

王 磊, 王洪彪, 马晓蕊, 钟 瑶, 赵继春 (59)

动物黑色素形成及调控机制研究进展

赵冲凯, 李丰敏, 吕腾飞, 王 坤, 黄 英, 贾俊峰 (65)

奶牛冷应激研究进展

宋佳星, 李尚法, 孙雨坤, 张永根 (72)

猪肠道菌群调控营养物质代谢的研究进展

武真邑, 刘 辉, 姜士良, 张慧燕 (77)

多组学生物信息学分析在家畜脂肪沉积研究中的应用研究进展

李 睿, 蔡 瑞, 连卫军 (83)

烟酰胺在猪营养中的应用研究进展

汤文杰, 何 鹏, 李书伟, 刁 慧, 晏军友, 江声耀 (91)

基于生物节律评价不同养殖工艺和饲养管理技术的动物福利和健康养殖问题

祝梦琦, 郭 琛, 王强军, 张利荣, 刘中英, 吴中红 (96)

褪黑素在绒山羊生产中的应用研究进展

姚凌云, 刁小东, 张 敏 (103)

瘦乳素及血管抑制素对哺乳动物卵泡发育的调控作用研究进展

谢社凤, 高凤茹, 李 莉, 蔡月贝, 姜永德, 卫恒习 (107)

羊毛性状相关蛋白、信号分子及基因的研究进展

冯梦雨, 姬磊磊, 伍修强, 李国富, 杨 果 (114)

发酵莱籽粕替代豆粕在动物生产中的应用研究进展

付晓晓, 曹维达, 谭品坤, 曹云鹏 (125)

科学技术

遗传育种

天衣麻鸡脂肪沉积候选基因表达水平与腹脂率的相关性分析

李群坤, 曹建朝, 罗超群, 谢 莉, 关林洪, 李 华, 向 海 (130)

基于 SNP 分子标记的云上黑山羊 3 个核心群遗传多样性分析

朱 兰, 王 鹏, 欧阳红兵, 江文彪, 兰 蕊 (136)

初产月龄对奶牛泌乳性能及长寿性的影响分析

董明明, 张梦华, 赵春香, 葛建军, 赵俊亮, 王 丹, 齐 磊, 钟雨伟, 黄杨霞, 王雅泰 (144)

鲁莱黑猪线粒体基因突变鉴别及其与生长和肉质性状关联分析

徐 磊, 蔡晋成, 郭仰诗, 徐廷强, 周孝生 (150)

关中奶山羊与阿尔卑斯奶山羊杂交后代乳品质研究

倪 洁, 胡家寿, 任俊杰, 曾 明, 梁旭辉, 谢建安 (157)

广灵大尾羊 *FITM2* 基因扩增、序列及表达特性分析

蔡国云, 乔利菊, 刘建华, 潘泽洋, 侯文斌, 李雨彤, 赵 祥, 刘文忠 (163)

副产蛋季产蛋情况及影响其蛋重初生重的因素分析

王 江, 阿卜杜热合曼·哈力克, 张梦华, 吴 飞, 何新宇, 石 兰, 马瑞兰, 张明, 苏来曼 (171)

祁县红牛 Y 染色体基因组遗传多样性与父系起源研究

雷振强, 王世强, 王开河, 夏小峰, 曹永霖, 雷初朝 (175)

犊牛生长发育对成年奶牛高峰奶产量和峰值日的影响

徐 明, 吴春霞, 王瑞宇, 陈 阮, 王 典 (178)

华东地区市场热点品种孵化与生产性能比较分析	张 楠, 金志明, 包 强, 李继杰, 张 勇, 孔令武, 陈伟明, 张 楠, 徐 强, 陈国宝 (186)
中国草原红牛 <i>PDK4</i> 基因克隆及其组织表达差异性研究	马彦茹, 于永生, 肖 威, 曹 阳, 姜立红, 赵玉民, 吴 健 (191)
牦牛附睾 <i>DPP4</i> 基因 <i>CDS</i> 区的克隆和表达研究	王红梅, 杨利娟, 吴德福, 杨健康, 郭旭明, 赵永生 (197)
阳原驴生长发育曲线拟合与分析	李 伟, 马晓君, 张同玉, 程 伟, 付永智, 苏咏梅, 朱文进 (202)
繁殖生理	
犏牛断奶前后血液生化指标和肠道通透性的变化规律研究	陈淑霞, 薛 林, 郭力明, 刁其玉, 李胜利, 毕明亮, 徐晓峰 (206)
褪黑素埋植剂对配种准备期水貂卵巢形态学和卵泡发育的影响	晏子超, 彭 静, 吴占斌, 高 志, 胡良臣, 马泽芳, 侯 凯 (211)
生物技术	
山羊 <i>NR1D1</i> 基因真核表达载体的构建及生物信息学分析	王选群, 高登科, 赵弘法, 李 丹, 马自荣, 刘静, 郭振刚, 靳亚平, 陈华涛 (215)
鸡 <i>IFITM3</i> 基因编码区克隆及生物信息学分析	张 贝, 孟 斌, 周泽宇, 张少军, 郭传生, 耿立斌, 李祥光 (223)
牦牛 <i>CST3</i> 基因的克隆、表达及生物信息学分析	李春奇, 陈晓英, 王红梅, 杨利娟, 赵永生 (230)
营养饲料	
蒲公英对肉仔鸡生长性能、营养物质表现代谢率及消化酶活性的影响	毛金菊, 段 琦, 郭 彦, 郭 彦, 尹 明, 王如心, 王 国, 安晓萍, 齐景伟 (235)
基于体外法评估稻草尼菜混贮饲料与精料的组合效应	李 翔, 陈 宁, 张振斌, 陈德飞, 于 刚, 王梦芝 (241)
椰子油对海南黑山羊羔羊屠宰性能及肉品质的影响	魏孟娟, 张 雨, 吴见霞, 刘 强, 施力光 (248)
丁香和槐角提取物对三黄鸡生长性能、抗氧化性能和脂质代谢的影响	张雪萍, 梁光智, 刁其玉, 吴智超, 崔明根 (253)
不同母体培养物对泌乳早期荷斯坦奶牛生产性能和瘤胃发酵参数的影响	许西峰, 王姗姗, 李 妍, 沈宜利, 王忠达, 任 帅, 杨金洋, 李建国, 高艳霞, 曹玉凤 (258)
小黑麦干草和羊草对德州驴的营养价值及其组合效应	陈永广, 董博超, 姜显新, 马献刚, 冯玉光, 曲清庭, 刘利林 (264)
复合营养素对羔羊生长性能、屠宰性能、肉品质及肌肉抗氧化能力的影响	郭 彦, 胡宇超, 杨 毅, 杨 磊, 郭志伟, 段 琦, 毛金菊, 王 国, 安晓萍, 齐景伟 (268)
黄花草秸秆的营养价值及其对绵羊瘤胃发酵参数的影响	田光元, 邵小杰, 焦 伟, 武喜明, 张洪霞, 张洪亮, 张建新 (276)
益生菌发酵甘草对甘草酸含量及其抑菌活性的影响	魏梅柳, 王留荣, 张同珍, 陈玲, 孟 洁, 宁尚楷, 李 伟 (282)
中产奶牛产前饲喂阴离子盐日粮对产乳性能和健康状况的影响	张 燕, 王倩倩, 孙晓梅, 杨占涛, 王中汉, 郝阳彪, 郭 成, 李胜利 (286)
脱氢乙酸钠对育成期水貂生长性能、营养物质表现消化率及脏器指数的影响	尹红古, 赵伟臣, 郝健中, 慈 洋, 李文生 (291)
复合益生菌固体发酵黑水虻幼虫过程中的成分分析	魏红芳, 刘 范, 赵向波, 唐桂蓉, 李正英, 唐春雨, 刘相慧, 郭建来 (296)
缓释尿素替代不同比例豆粕对瘤胃体外发酵参数的影响	王里河, 纪守坤, 李金辉, 陈佳欣, 田立霞, 杨彩虹, 张立中, 段春辉, 刘月琴, 严 昱, 张英杰 (300)
活性干酵母对高温应激条件下母猪体况、繁殖性能、乳成分和血清生化指标的影响	任 莹, 关泽涛, 周世龙, 赵元琪, 张 晨, 吕良康, 雷 亮, 徐学文, 赵胜军 (304)
不同胎次的干奶期奶牛瘤胃温室气体排放量特征及对泌乳性能的影响	刘忠豪, 董利峰, 董 鹏, 杨 琦, 高厚华, 田志红, 吴秋红, 刁其玉 (309)
生产与管理	
超微粉催情散提高产后奶牛发情受孕效果的研究	胡婷婷, 王翠柳, 余斯亮, 陈良雷, 王 峰, 田健亮, 郭凯军 (315)
类芽孢杆菌产脂环七肽的中试发酵参数优化	黄 升, 陈 利, 苏国银, 黄金奇 (319)

产业经济

产业透视

2021 年我国驴业发展概况、现实困境与突围之路

袁露露, 魏德飞, 刘德德, 王春洪, 王玉斌 (323)

行业调研

非洲猪瘟病毒在饲料中的传播风险及缓解措施

李 鹏, 胡玉云, 孙得发, 周 阳 (328)

目次

10
2022

第 58 卷 第 10 期

月刊

支持单位

四川省畜牧饲料有限公司

Chinese Journal of Animal Science

Vol. 58, No. 10, 2022

Selected Contents

Current Knowledge and Application of Runs of Homozygosity in the Commercial Lean Pig Genome	LI Guixin, WANG Shiyuan, YANG Jie, WANG Xiaojing, ZHENG Enqin (1)
Advances in Ruminal Biohydrogenation of Unsaturated Fatty Acids and Its Bio-Function on Lactating Performance in Dairy Cattle	PU Xianxun, LI Qingshuang, WANG Min, TAN Zhiliang (8)
Advances in the Antioxidant Function and Application of Curcumin in Animal Production	FU Jiangwan, ZHANG Haibo, OUYANG Hongbing, LI Lizhi, LIAO Xiaoping, GUAN Wukun, GUO Dongsheng (14)
Research Progress on n-3 Polyunsaturated Fatty Acid Deposition in Lamb Meat	ZHANG Zhanyang, ZHANG Boyan, LUO Hailing, WANG Bing (19)
Research Progress on the Application of Low Protein Diet with Balanced Amino Acids and Its Effect on Pork Quality	YANG Jingsen, WANG Li, JIANG Zongyong, ZHU Cui (24)
Research Progress on Oxidative Stress Biomarkers in Different Body Fluids of Dairy Cows	LUO Lin, CAI Jie, DENG Mengmeng, LIU Jianxin, WANG Dingming (31)
Preparation and Evaluation of Coated Feed Additives: A Review	WU Peng, XU Erhua, ZHAO Wei, HOU Jia, ZHANG Qianyan, WANG Xiao, LI Miao, FAN Haiping, CHENG Saisai, LI Zhefeng (37)
New Progress on Single Cell Sequencing Technology and Its Application to Animal Genetics and Breeding	PENG Wei, JIA Kexin, ZHANG Zijing, LIU Xian, CAI Cuicui, LI Xintiao, LEI Chuzhao, CHEN Hong, HUANG Yongzhen (44)
The Research Progress in Methylation Modification of m6A and its Molecular Mechanism Affecting the adipogenesis of Animals	SONG Xingyu, PENG Wei, LIU Xian, YAO Zhi, ZHANG Zijing, LEI Chuzhao, WANG Eryao, HUANG Yongzhen (52)
Research Progress on Extract and Physiological Function of Curcumin and Its Application in Broiler Production	WANG Min, WANG Hongbao, MA Xiaoni, ZHONG Yao, ZHAO Jichun (59)
Research Progress on the Formation and Regulation Mechanism of Animal Melanin	ZHAO Jingying, LI Fengyun, DOU Tangfei, WANG Kun, HUANG Ying, JIA Junjing (65)
Research Progress on Cold Stress in Dairy Cows	SONG Jiamin, LI Shangru, SUN Yukun, ZHANG Yonggen (72)
Research Progress on Regulation of Nutrient Metabolism by Intestinal Microbiota in Pigs	WU Zhenyi, LIU Hui, QIN Shizhen, ZHANG Dongyuan (77)
Application and Research Progress in Multi-Omics Bioinformatics Analysis in the Fat Deposition in Livestock	LI Rui, CAI Rui, PANG Weiqun (83)
Research Progress on the Application of Valine in Swine Nutrition	TANG Wenjie, HE Peng, LI Shaoqi, DIAO Hai, YAN Jiyun, KUANG Shengyao (91)
Evaluation of Animal Welfare and Healthy Breeding under Different Raising Processes and Feeding Management Techniques Based on Biological Rhythms	ZHU Mengqi, GUO Yao, WANG Qiangjun, ZHANG Kehai, LIU Zhongying, WU Zhonghong (96)
Advances in the Application of Melatonin in Cashmere Goat Production	YAO Liangyun, DIAO Xiaoguo, ZHANG Wei (103)
Research Progress on the Regulation of Prolactin and Vasointerin on Follicular Development in Mammals	XIE Shifeng, GAO Fenglei, LI Ju, HAN Beibei, LI Chengde, WEI Hengxi (107)
Research Progresses of the Related Proteins, Signal Molecules and Genes to Wool Traits	FENG Mengyu, JI Kanxi, WU Xukun, XIN Guosheng, YANG Guo (114)
Research Progress on Application of Replacing Soybean Meal with Fermented Rapeseed Meal in Animal Production	FU Xiaopi, SHANG Zhenda, TAN Zhankun, CAO Yunhe (125)
Correlation Analysis of Adipose Deposition Candidate Gene Expression Level and Abdominal Fat Percentage in Tianmang Partridge Chicken	LI Xiangkun, ZENG Dianshu, LUO Chuawei, XIE Li, WU Linxi, LI Hua, XIANG Hai (130)
Genetic Diversity Analysis of 3 Core Yunshang Black Goat Populations based on SNP Markers	ZHU Lan, WANG Peng, OUYANG Yina, JIANG Yanting, LAN Rong (136)
Search for Mutations of Mitochondrial Gene and Association Analysis with Growth and Meat Quality Traits in Lushan Black Pig Population	XU Jing, ZHANG Changzheng, ZHENG Yangqing, ZHANG Tingrong, ZHOU Lisheng (150)
Research on the Law of Change in Blood Biochemical Indexes and Intestinal Permeability of Calves before and after Weaning	CHEN Zhihui, CHEN Lin, ZHANG Lili, DIAO Qiye, LI Shengli, BI Yanliang, XU Xiaofeng (166)
Construction of a Eukaryotic Expression Vector of Goat NR1D1 Gene and its Bioinformatics Analysis	WANG Yiqun, GAO Dangle, ZHAO Hongcong, LI Dan, MA Baorong, LIU Pangao, SUN Yuetong, JIN Yaping, CHEN Hantao (215)
Effects of Dandelion on Growth Performance, Nutrient Apparent Metabolic Rate and Digestive Enzyme Activity of Broilers	MAO Jinjie, DUAN Ting, GUO Tao, GUO Hui, YIN Na, WANG Ruxin, WANG Yuan, AN Xiaoping, QI Jingwei (235)
Evaluation of Combined Effects of Concentration and Mixed Silage with Rice Straw and Chinese Cabbage Waste <i>in vitro</i>	LI Chuang, CHEN Ning, ZHANG Zhenbiao, CHEN Yifei, YU Xiang, WANG Mengzhi (241)
Effects of Coconut Oil Supplementation on Carcass Traits and Meat Characteristics in Hainan Black Goat Lambs	Ji Mengyao, ZHANG Yu, WU Lingli, LIU Qiang, SHI Liguang (248)
Effects of Compound Nutrients on Growth Performance, Slaughter Performance, Meat Quality and Antioxidant Index in Muscle of Lambs	GUO Tao, HU Yuchao, YANG Yi, YANG Lei, CUI Zhiwei, DUAN Ting, MAO Jinjie, WANG Yuan, AN Xiaoping, QI Jingwei (268)
Transmission Risk and Mitigations of African Swine Fever Virus in Feed	LI Peng, MU Yuyun, SUN Defa, ZHOU Yang (328)

69

1953-2022



微信扫描二维码

承一脉学术为根本，秉一颗专注做内容。
集一万平台展硕果，写一段故事见真章。

《中国畜牧杂志》创刊于1953年，是中国科学技术协会主管、中国畜牧兽医学会主办、中国农业大学为学术支撑单位的学术期刊。期刊现已被全国中文核心期刊数据库、中国科技核心期刊数据库、中国学术期刊综合评价数据库、中国核心期刊（遴选）数据库、维普网等多家数据库收录。

根据《中国学术期刊影响因子年报（自然科学与工程技术·2021版）》数据，《中国畜牧杂志》影响力指数（CI值）提高到613.37，学科排序为7/70，在畜牧类杂志中排名第3位，复合影响因子达到1.25；在《中国学术期刊评价报告（第六版）》中，《中国畜牧杂志》获“权威学术期刊（A+）”评价，在全国“畜牧、兽医科学”类90本期刊中排名第三。

《中国畜牧杂志》始终坚持以“内容为王”为宗旨，秉承客观、公正、科学的传播理念，以引导正确的科技方向为己任，近年来重点打造了综述、科学技术、产业经济、特别报道等传统栏目和遗传育种、繁殖生理、饲料营养、检测技术、食品安全、环境控制与动物福利、动物健康、科技应用、产业透视、畜牧经济等二级栏目，在畜牧、兽医科学类期刊中具有鲜明特色。

ISSN 0258-7033
CN11-2083/S

国内邮发代号：82-147

国外发行代号：SM452

定价：50.00元



微信订阅

瘦肉型猪基因组中 ROH 的应用及研究进展

李桂新, 王诗媛, 杨杰, 王小鹏*, 郑恩琴*

(华南农业大学动物科学学院, 国家生猪种业工程技术研究中心, 广东广州 510642)

摘要: 连续性纯合片段 (Runs of Homozygosity, ROH) 是研究群体历史和性状特征的窗口, 被广泛用于检测动物全基因组近交水平和识别基因组选择区域。目前, 大量单核苷酸多态性 (Single Nucleotide Polymorphisms, SNP) 基因型的可用性使得基因组中 ROH 的特征变得更加准确, 但缺乏可靠和统一的 ROH 分析指南。本文介绍了 ROH 发展简史、检测软件及参数, 综述了 ROH 在杜洛克、大白、长白、皮特兰猪等商业瘦肉型猪遗传育种中的应用, 包括遗传多样性的评估、近交系数的度量、人工选择痕迹的识别和不利 ROH 的检测, 并对 ROH 在猪育种应用中的发展前景进行了展望。

关键词: 瘦肉型猪; ROH; 群体历史; 人工选择; 不利 ROH

中图分类号: S828.2

文献标识码: A

DOI 编号: 10.19556/j.0258-7033.20210919-01

在二倍体生物的基因组中, 连续性纯合片段 (Runs of Homozygosity, ROH) 是指基因组中没有杂合子的纯合基因型的连续染色体片段^[1]。ROH 主要由群体历史变迁形成的, 如种群瓶颈、遗传漂变和近亲繁殖等^[2]。另外, 强烈的自然和人工选择也塑造着基因组区域不同的 ROH 模式^[3]。较长的单倍型遗传自最近的共同祖先, 较短的单倍型遗传自遥远的祖先^[4]。ROH 被认为是评估人类和动植物近亲繁殖的有效方法^[5]。优良性状的选择产生了畜禽品种显著的表型变化, 并形成了基因组不同区域的 ROH 模式^[4,6]。同时, 选择也增加了基因座周围的纯合性。育种计划的选择强度和优良种畜的广泛应用导致近亲繁殖增加, 引起群体的遗传多样性降低。使用中、高密度 SNP 基因芯片扫描基因组中的 ROH 是识别血缘同源 (Identical By Descent, IBD) 单倍型的有

效方法。基因组 ROH 可以提供一个种群过去和最近的群体变化信息, 揭示世代间演变的历程^[7-8]。最近几十年, 人类对瘦肉型猪的高强度选育可能导致近亲繁殖的累积, 造成基因组中单倍型的多样性降低, 等位基因的基因频率升高甚至接近固定存在于基因组中。本文主要围绕动物基因组 ROH 研究的发展简史、ROH 的识别软件和相关参数, 以及 ROH 在瘦肉型猪的应用和研究进展进行综述, 旨在为瘦肉型猪的育种提供参考依据。

1 ROH 发展简史

1999 年, Broman 等^[9]首次对人类基因组进行连续性纯合检测分析, 推断纯合片段可能代表同源性, 而且可能与人类健康具有很大的相关性。2006 年, Gibson 等^[1]利用高密度 SNP 芯片首次在人类基因组中报道了 ROH, 揭示了整个基因组中 ROH 的长度、频率和分布。2010 年, Sölkner 等^[10]、Ferenčaković 等^[11]首次将高密度 SNP 芯片应用于畜禽基因组 ROH 研究。随着重测序技术和 SNP 芯片的发展、成熟和广泛引用及成本不断降低, 出现了一系列关于畜禽基因组 ROH 分析的相关研究。例如在猪基因组研究中, Bosse 等^[4]2012 年, 首次研究瘦肉型猪基因组 ROH 的群体历史 (如群体瓶颈、近交繁殖等)、基因组特征 (如 GC 含量和重组率) 和选择对基因组 ROH 的影响。Silió 等^[12]利用系谱和基因组数据中评估了伊比利亚猪猪的近交水平和近交

收稿日期: 2021-09-19; 修回日期: 2021-11-12

资助项目: 广东省现代农业产业技术体系创新团队 (2021KJ26); 广东省级乡村振兴战略专项 (200-2018-XMZC-0001-107-0145); 广东省自然科学基金 (2018B030313011)

作者简介: 李桂新 (1995-), 男, 江西赣州人, 硕士, 主要从事分子遗传育种与繁殖研究, E-mail: guixinli2016@hotmail.com

* 通讯作者: 郑恩琴 (1981-), 女, 广东惠州人, 硕士, 高级实验师, 主要从事遗传育种研究, E-mail: eqzheng@scau.edu.cn; 王小鹏 (1988-), 湖北荆门人, 博士, 主要从事动物群体遗传及复杂性状解析, E-mail: xiaopengwang2017@hotmail.com

衰退。Herrero-Medrano 等^[13]利用高密度 SNP 芯片通过 ROH 分析,揭示了家猪和野猪的群体历史和近交繁殖。2014 年, Herrero-Medrano 等^[14]使用 60K SNP 芯片和重测序数据 (~10X) 进行 ROH 检测,发现检测的 ROH 和近交系数呈现很强的相关性。此后,瘦肉型猪基因组 ROH 成为一个兴起的研究热点。

2 ROH 检测的软件和参数

目前,对于不同的数据集,关于不同软件间以及同一软件间不同参数的设置对 ROH 检测影响的研究较少,使用不同软件之间的最佳参数难以确定。不同的研究采用了不同的软件和参数来识别 ROH。目前,缺乏对 ROH 的定义和识别标准,限制了 ROH 相关研究的发展。不同的研究在检测 ROH 时使用了各自的标准,有些研究采用了严格的标准,如 Lencz 等^[17]采用至少 100 个连续 SNPs,不允许存在杂合子,且需要在 10 个或者更多个体间共享的 ROH 为研究目标。而有些研究采用了较为宽松的标准,如 Spain 等^[15]使用完全连锁不平衡和低连锁不平衡的数据集,研究了不同数量的 SNPs 和不同长度的 ROH 对识别 ROH 的影响,研究允许 2% 的杂合子出现,且取消了对 SNPs 之间最大距离和最小 SNPs 密度的限制。定义和识别 ROH 标准的差异可能导致千差万别的结果,同时增加了假阳性出现的概率,并使不同研究结果之间的比较变得困难^[13]。

2.1 检测 ROH 的软件 当前,根据不同的方式识别基因组 ROH,可以将检测方法分为连续性 (Consecutive Runs) 识别和窗口滑动 (Sliding Window) 识别。连续性识别是一种沿着基因组对 SNPs 位点连续扫描检测的方法,其设置参数类别包括一个 ROH 中最小的 SNPs 数量、最大的杂合子与缺失 SNPs 数量、相邻 SNPs 位点的最大间隙以及最小运行长度等。应用连续性识别方法检测 ROH 的软件有 SVS (Golden Helix SNP & Variation Suite v.7.6.8)、SAS (SAS Institute) 以及最近比较热门的 R 包—detectRUNS^[16]等。另一种检测方法是窗口滑动检测法,即设置窗口大小,以滑动窗口形式检测基因组 ROH 的方法。常用软件及脚本有 PLINK^[17]、GERMLINE^[18]、BEAGLE^[19]、cgaTOH^[20]、SVS (Golden Helix SNP & Variation Suite v.7.6.8)、BCFtools/RoH^[21]、RZooRoH^[22]和 FORTRAN 脚本^[23]等。其中,SVS 软件和 R 包—detectRUNS 能同时用这 2 种

检测方法来识别基因组 ROH。目前 PLINK 的窗口滑动和 R 包—detectRUNS 的连续性检测方法被广泛应用于 ROH 的研究。

2.2 检测 ROH 的参数 无论是连续性运行还是窗口滑动识别 ROH,都需要对多个参数和阈值进行设置。这些设置可能会对 ROH 数量、大小和分布产生显著影响,而且默认设置值并不总是适合各种基因型数据,尤其是不同密度的商业 SNP 芯片数据。Purfield 等^[8]使用牛 50K 和 HD SNP 芯片 (777972 SNPs) 分析 ROH 时,发现 HD 芯片能更准确地识别 ROH,50K 芯片的 SNP 密度适用于大多数牛品种的 ROH 检测。作者还比较了 50K 和 HD 芯片共有 SNPs 数据集与 HD 芯片识别 ROH 的差异,发现二者都难以识别 0.5~1 Mb 短长度的 ROHs,且都能识别所有 >5 Mb 长度的 ROHs。同时,因为低密度 SNP 芯片估计的 ROH 水平与根据系谱的近交系数呈正相关。因此,常用商业基因芯片的标记密度在鉴定 ROH 是合适的。进行 ROH 分析之前,是否需要过滤最小等位基因型频率 (Minor Allele Frequency, MAF)、偏离哈代-温伯格平衡 (Hardy-Weinberg Equilibrium, HWE) 和高连锁不平衡 (Linkage Disequilibrium, LD) 的 SNPs 尚未得到共识。大多数研究都是采用过滤 MAF<0.01 或者 <0.05 的 SNPs。有些研究也过滤 LD,如 Meyermans 等^[24]测试过滤不同 LD 和 MAF 值对检测 ROH 的影响,发现随着过滤 LD 值的增加,基因组覆盖率迅速下降;同时发现 MAF 过滤不仅影响固定区域的 ROH 检测,而且在非固定区域检测到 ROH 的发生率也存在较大差异。因此,在进行 ROH 研究时,作者建议不要进行 LD 和 MAF 过滤。SNP 基因型分型错误是另一个可能影响 ROH 检测的因素,它的存在会影响包含 SNPs 数量多的长 ROH 识别。当前研究的一个解决方案是允许一定数量的杂合 SNP 存在,但这是否会影响检测 ROH 的可靠性尚未得到证实^[24]。此外,其他因素如 ROH 中 SNP 的最小密度、最大间隙、滑动窗口大小、窗口阈值、ROH 最小长度、ROH 包含的最少 SNP 数量和最多缺失基因型数量在一定程度上都会影响 ROH 的检测。

目前,ROH 分析广泛应用于瘦肉型猪研究中,主流瘦肉型猪品种基因组 ROH 的部分研究工作以及用于识别 ROH 的相应参数和阈值均在表 1 中展示。然而还有一些研究的参数设置没有明确地提到,不确定作者是

表 1 不同瘦肉型猪品种 ROH 识别和参数设置的比较

品种	数据集	软件	最少连续 SNP ^a	最小密度, SNP/Kb ^b	最大间隙, Kb ^c	最小长度, Kb ^d	允许杂合基因型数量 ^e	允许缺失基因型数量 ^f	滑动窗口大小 ^g	阈值 ^h	参考文献
杜洛克 / 汉普夏 / 大白 / 长白猪	60K	PLINK V1.07	20	1/1000	-	-	1	-	10	-	[4]
杜洛克 / 大白 / 白杜洛克 / 长白猪	60K	PLINK V1.07	50	-	-	500	1	5	50	-	[25]
杜洛克 / 大白 / 长白猪	60K/ 重测序	PLINK V1.07	10	1/1000	1000	10	-	-	100	-	[14]
大白 / 长白	60K	PLINK	50	-	-	1000	1	1	50	-	[26]
杜洛克 / 大白 / 长白 / 巴克夏 / 皮特兰猪	重测序	PLINK V1.90	-	1/100	250	1000	0	-	50	0.05	[27]
皮特兰猪	60K/80K	PLINK V1.90/ RZooRoH	53/57/55/ 47/49	1/150	1000	1000	0	0	-	0.05	[28]
杜洛克 / 大白 / 长白猪	80K	PLINK V1.90	-	-	-	500	1	5	50	-	[29]
大白 / 长白猪	60K	detectRUNS 包	43	1/500	1000	1000	1	2	50	0.05	[30]
杜洛克 / 大白 / 长白猪	1.4M	PLINK V1.90	-	-	-	500	1	5	50	-	[31]
杜洛克 / 大白 / 长白猪	60K	PLINK V1.90	15	1/1000	1000	1000	0	-	-	-	[32]
皮特兰猪	50K	PLINK V1.07	10	-	1000	1000	1	1	-	-	[33]
长白猪	60K	cgaTOH	30	-	1000	-	0	-	-	-	[34]
杜洛克 / 大白 / 长白猪	80K	PLINK V1.90	80	1/100	1000	1000	1	5	50	-	[35]
大白猪	50K	PLINK V1.90	50	1/100	1000	1000	1	5	50	0.01	[36]
杜洛克 / 大白 / 长白 / 巴克夏 / 皮特兰猪	60K	PLINK V1.90	-	1/150	1000	-	0	1	-	0.05	[37]

注: a、b、c、d、e 和 f 分别表示一个 ROH 中允许的最少连续的 ROH 数量、最小密度、最大间隙、定义的最小长度、存在的最多杂合基因型数量和最多缺失基因型数量; g 表示由 SNPs 窗口滑动; h 表示假阳性 ROHs 的百分比; “-” 表示文献没有提供相关信息。

否使用默认设置或调整参数, 这样可能使得基因组的覆盖率被过高或过低估计。可见, ROH 的识别和定义标准在不同品种内和品种间都存在差异。

3 ROH 在瘦肉型猪基因组的应用

2010 年, Sölkner 等^[10]首次在牛基因组研究中报道 ROH 后, 越来越多学者对畜禽的 ROH 进行了更深入的探索, 尤其是瘦肉型猪 ROH 的研究, 包括遗传多样性分析、群体历史及近交系数的评估、选择信号的鉴别和不利 ROH 检测等多个重点研究方向。

3.1 ROH 分析遗传多样性 基因组信息已被广泛用于评估畜禽的遗传多样性。经历过高强度人工选育的优良畜禽动物的表型变异减少, 基因组单倍型的多样性降低, 被选择位点周围的纯合性增加, 出现 ROH 模式^[38], 导致优良动物的遗传多样性降低、近交系数升高。

现在, 越来越多研究将 ROH 作为分析群体遗传多样性的指标, 或者将西方商业猪的 ROH 当作地方猪的背景数据进行比较^[25,29,31,35]。Ai 等^[25]观察到西方瘦肉型猪种(白色杜洛克、杜洛克、大白和长白猪)比中国地方猪种表现出更长的 ROH, 这与瘦肉型猪经历过高

强度选育使得基因组纯合位点增多, 导致遗传多样性降低的理论是一致的。在瘦肉型猪中, 白色杜洛克猪的 ROH 是最低的, 但 $r_{0.3}^2$ 值是最高的, 这可能是因为白色杜洛克猪是由杜洛克和大白猪培育而来, 最近的杂交事件导致了长距离的 LD 范围出现, 同时减少了基因组中 ROH 的长度。Zhang 等^[27]同样比较了商业瘦肉型猪和中国地方猪基因组的 ROH 模式, 瘦肉型猪比中国地方猪具有更多小于 1 Mb 的 ROH, 瘦肉型猪 ROH 的平均长度普遍大于中国地方猪。因此, 与中国地方猪相比, 瘦肉型猪的遗传多样性可能会进一步降低。这些结果显示了每个猪种的平均 ROH 数量和不同的平均基因组覆盖率。然而目前的研究结果并不多, 很难直接比较不同品种或者同一种群间的 ROH 模式。此外, 需要判断研究结果是否真实反映了实际情况, 因为不同的 ROH 定义标准和设置参数偏差可能会导致过高或过低的 ROH 估计。

3.2 ROH 评估近亲繁殖水平 近交系数(Inbreeding Coefficient, F)是监测种群遗传变异和管理畜禽遗传资源的重要参数之一, 准确地评估个体和群体的近交系数不仅是实际生产的切实需求, 也是近交效应研究的重

点。传统上,近交系数的估计是基于系谱信息(F_{PED}),该方法度量的是个体任意位点上出现共享等位基因型的概率,估计值是固定的^[39]。 F_{PED} 值是否能够真实反映群体的近交水平取决于个体系谱信息的完整性和准确性。然而,在实际生产中,由于各种因素导致系谱信息不完整或者记录错误的现象普遍存在,这将大大降低 F_{PED} 估计值的可靠性。此外, F_{PED} 依赖于基础群信息,基于群体的基因组没有经历重组和个体不受选择的基础计算,也难以反映染色体中的纯合性,更不允许计算特定染色体区域的近交系数。

早期的低密度分子标记由于标记数量有限,可能存在抽样误差,导致估计值的准确性低。但随着重测序技术和高密度 SNP 芯片的发展, DNA 分子信息显著地提高了评估近交系数的准确性。利用分子信息来评估群体的近交水平主要分为 3 种方法:基于纯合子 SNP 的近交系数(F_{HOM});基于 SNP 分子标记间的近交系数(F_{GRM});基于 ROH 的近交系数(F_{ROH})。与 F_{HOM} 和 F_{GRM} 相比, F_{ROH} 估计群体近交系数有以下几点优势:第一,可以区分是 IBD 还是状态同源 (Identical By State, IBS)。在使用单一分子标记计算估计近交水平时,难以区分 IBD 和 IBS,但使用 ROH 却可以做到;第二,可以揭示近交发生的时期(近期还是遥远时期)。在减数分裂时重组事件可以打断较长的 ROH 片段,因此 ROH 的长度和距离共同祖先的世代数有很强的相关性。例如 Shi 等^[36]检测到大白猪的近交水平可能是在近 5 代积累导致,而且 F_{ROH} 和 F_{PED} 之间的相关性比较低(0.18~0.37)。这个结果与其它瘦肉型猪的研究一致,并且 F_{ROH} 和 F_{PED} 之间的相关性在不同长度的 ROH 没有明显变化。 F_{HOM} 不能区分 IBD 和 IBS,可能过高地估计了近交水平,除了与 F_{GRM} 之间的相关性较低外,与 F_{PED} 和不同长度的 F_{ROH} 相关性都较高,这在皮特兰猪、杜洛克猪、长白猪以及不同品系大白猪的研究中都得到相同的结论^[26,30,32-33,40]。Zhan 等^[33]基于系谱信息和基因组信息评估了皮特兰猪的近交水平, F_{HOM} 和 F_{ROH} 有显著的高相关性($r=0.949$), F_{ROH} 与 F_{PED} 之间相关性较低,这也与 Gorssen 等^[37]在其他皮特兰猪群体的报道一致, F_{HOM} 和 F_{ROH} 都与 F_{PED} 没有高相关性。这些结果表明 F_{ROH} 和 F_{PED} 之间的相关性不高,可能是因为 F_{PED} 和 F_{ROH} 应用的计算方法和原理不同,如 F_{PED} 不能计算基因组中 IBD 实际比例,也不能解释减数分

裂期间重组的随机事件^[41-42],这可能影响了计算的准确性,特别是没有或缺乏系谱信息时^[33]。同时基因组中某些单倍型可能是由于局部的低程度重组和高水平 LD 形成的^[43]。此外,相关研究表明短片段和中等长度片段的 ROH 更有可能来自 LD 或者遥远时期的近交事件,而通过排除 LD 和随机效应的干扰,长片段 ROH 估算的近交系数可以更准确地代表近期近交繁殖的水平^[44]。然而,不同群体的 ROH 分类没有固定标准。因此,使用总长度的 ROH 来估算近交水平是一个更为可行的方法^[45-46]。由评估瘦肉型猪和其他动物 ROH 的相关研究表明, F_{ROH} 可能是一种在理论上更有效和准确的度量畜禽近交水平的替代方法,被广泛用于评估个体和群体的近交水平^[3]。

3.3 ROH 识别人工选择的痕迹 欧洲家猪大约在 9 000 年前在近东被驯化,是最早被驯化的家畜之一^[47-48]。近一两百年以来,随着人类对瘦肉型猪品种的需求和优良种畜的选育,不仅造成了瘦肉型猪外形和生长性状的变化,还在基因组中留下了选择的印迹^[49]。当基因组的特定区域受到长期、连续和高强度的正向选择,特别是人工选择,会表现出受选择区域的纯合频率迅速增加,甚至固定下来,出现 ROH 片段。Bosse 等^[4]观察到, ROH 在染色体中是非随机分布的,许多 SNPs 分子标记在 ROH 片段中具有异常频率,这些区域被称为“ROH 热点”或“ROH 岛”。因此,通过扫描基因组中的 ROH 高频区域区段可以鉴定受到选择的基因组区域。

Gorssen 等^[28]对 5 个不同品系皮特兰猪基因组进行 ROH 分析,并与杜洛克猪、大白猪和长白猪进行比较,在 8 号染色体上观察到一个 90 Mb 的 ROH 热点区域,约 85% 的个体存在共享 ROH,著名的影响猪毛色的 *KIT* 基因在这个区域也被鉴定到。此外, 50~60 Mb 区域处发现所有皮特兰品系、长白猪和大白猪存在的 ROH 岛重叠。Zhan 等^[33]同样进行了皮特兰猪 ROH 分析,将 16 个 ROH 热点区域与猪数量性状基因座 (Quantitative Trait Locus, QTL) 进行比对,发现其与生长、肉质、胴体品质和繁殖等性状相关。同时,还鉴定到在生物学过程起关键作用的候选基因。Shi 等^[36]在大白猪基因组的 4 个 ROH 热点区域中检测到大量与繁殖相关的候选基因,这与大白猪繁殖力高、产仔数多的种质特性相一致。Gorssen 等^[37]利用全球共享的动物基因组数据进行了 ROH 分析,观察到杜洛克猪在这

些动物中 ROH 热点区域最多, 同时 SNPs 发生频率也是最高的, 很大部分超过 80%, 有些甚至达到 100%。

经过长期的正向选择, 控制重要经济性状的调控区域将高度纯合, 这些特定区域也就更容易暴露在 ROH 高频区域。所以, 观察到的 ROH 高频区域有助于识别被选择的位点或区域, 寻找与经济性状相关的候选基因, 并应用于育种实践中。

3.4 检测不利 ROH 单倍型 在瘦肉型猪育种工作中, 重要经济性状的遗传增益是通过高强度选育实现的^[50]。然而, 高强度选育会造成遗传多样性降低和近亲繁殖程度增加, 而且可能会导致隐性有害等位基因的积累。这反过来又导致近交衰退, 即平均表型性能的降低^[51]。近亲繁殖会产生 ROH, 大多数的 ROH 对表型有中性甚至有利的影响, 从而掩盖了少数不利的 ROH 影响^[52]。由于近亲繁殖的模式在整个基因组中是不同的, 因此, 确定与表型相关的基因组 ROH 区域可以使遗传增益和近亲繁殖水平之间得到更有效的平衡^[50]。有不少研究报道了 ROHs 对经济性状会造成不利影响^[53-55], 但是没有考虑到不同 ROH 片段对同一表型的影响和不同 ROH 片段对多个表型的影响。Howard 等^[56]提出了一种能检测对表型不利的 ROH, 即能对个体内和个体间 ROH 出现的联合效应进行描述的软件——Unfavorable Haplotype Finder。利用此软件在大白猪基因组中检测到 13 个生长性状和繁殖性状相关的不利 ROH (平均长度为 1.54 Mb; 长白猪中为 4 个 ROHs, 平均长度为 1.56 Mb), 以及 2 个品种中 3 个与繁殖性状相关的共有区域。同时还指出 LD、QTL 和 ROH 在基因组中发生的频率对识别不利 ROH 的能力产生很大影响。

4 小结与展望

ROH 分析是研究畜禽基因组特征的有效方法。本文针对瘦肉型猪基因组中 ROH 在分析遗传多样性、评估近交繁殖水平和识别人工选择痕迹等应用进行了全面的综述。然而, 在瘦肉型猪中有害突变的累积是否在 ROH 出现的概率更高, 以及 ROH 是否与经济性状具有强相关性尚未有报道, 仍需要进一步探究。

虽然目前对基因组中 ROH 有很多研究, 但仍存在许多科学问题值得进一步挖掘: ①瘦肉型猪中 ROH 的定义以及检测的参数和阈值缺乏一致的标准, 很难直接比较不同研究的结果; ②瘦肉型猪 ROH 覆盖率高的

ROH 区域是否存在有害突变, 不同长度的 ROH 中有害突变的频率是否有规律, 对生长性状是否产生负面影响; ③ Unfavorable Haplotype Finder 软件既然能找出不利 ROH 片段, 那么是否可以能通过调整参数找到对瘦肉型猪最有利的 ROH 片段或者杂合子片段 (Runs of Heterozygosity, ROHet) 以及不利的 ROHet 片段, 以用于现代育种计划。因此, 关于瘦肉型猪基因组中 ROH 和 ROHet 仍需深入研究, 进而为育种工作和解析瘦肉型猪遗传和表型差异的遗传机制提供理论依据。

参考文献:

- [1] Gibson J, Morton N E, Collins A. Extended tracts of homozygosity in outbred human populations[J]. *Hum Mol Genet*, 2006, 15(5): 789-795.
- [2] Ceballos F C, Joshi P K, Clark D W, et al. Runs of homozygosity: windows into population history and trait architecture[J]. *Nat Rev Genet*, 2018, 19(4): 220-234.
- [3] Peripolli E, Munari D P, Silva M V G B, et al. Runs of homozygosity: current knowledge and applications in livestock[J]. *Anim Genet*, 2017, 48(3): 255-271.
- [4] Bosse M, Megens H, Madsen O, et al. Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape[J]. *PLoS Genet*, 2012, 8(11): e1003100.
- [5] Keller M C, Visscher P M, Goddard M E. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data[J]. *Genetics*, 2011, 189(1): 237-249.
- [6] Nothnagel M, Lu T T, Kayser M, et al. Genomic and geographic distribution of SNP-defined Runs of homozygosity in Europeans [J]. *Hum Mol Genet*, 2010, 19(15): 2927-2935.
- [7] Lencz T, Lambert C, Derosse P, et al. Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia[J]. *Proc Natl Acad Sci U S A*, 2007, 104(50): 19942-19947.
- [8] Purfield D C, Berry D P, Mcparland S, et al. Runs of homozygosity and population history in cattle[J]. *BMC Genet*, 2012, 13(1): 70.
- [9] Broman K W, Weber J L. Long homozygous chromosomal segments in reference families from the centre d'Etude du polymorphisme humain[J]. *Am J Hum Genet*, 1999, 65(6): 1493-1500.
- [10] Sölkner J, Ferencakovic M, Gredler-Grandl B, et al. Genomic metrics of individual autozygosity, applied to a cattle population [Z]. *Heraklion*, 2010.
- [11] Ferencakovic M, Hamzic E, Gredler B, et al. Runs of homozygosity reveal genome-wide autozygosity in the austrian fleckvieh cattle[J]. *Agriculturae Conspectus Scientificus*, 2011, 76(4): 325.
- [12] Silió L, Rodríguez M C, Fernández A, et al. Measuring inbreeding and inbreeding depression on pig growth from pedigree or SNP-derived metrics[J]. *J Anim Breed Genet*, 2013, 130(5): 349-360.
- [13] Herrero-Medrano J, Megens H, Groenen M, et al. Conservation

- genomic analysis of domestic and wild pig populations from the Iberian Peninsula[J]. *BMC Genet*, 2013, 14: 106.
- [14] Herrero-Medrano J, Megens H, Groenen M, *et al*. Whole-genome sequence analysis reveals differences in population management and selection of European low-input pig breeds[J]. *BMC Genomics*, 2014, 15: 601.
- [15] Spain S L, Cazier J, Houlston R, *et al*. Colorectal cancer risk is not associated with increased levels of homozygosity in a population from the United Kingdom[J]. *Cancer Res*, 2009, 69(18): 7422-7429.
- [16] Filippo B, Paolo C, Giustino G, *et al*. detectRUNS: Detect Runs of homozygosity and Runs of Heterozygosity in diploid genomes[DB/CD]. 2019.
- [17] Purcell S, Neale B, Todd-Brown K, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses[J]. *Am J Hum Genet*, 2007, 81(3): 559-575.
- [18] Gusev A, Lowe J K, Stoffel M, *et al*. Whole population, genome-wide mapping of hidden relatedness[J]. *Genome Res*, 2008, 19(2): 318-326.
- [19] Browning S R, Browning B L. High-resolution detection of identity by descent in unrelated individuals[J]. *Am J Hum Genet*, 2010, 86(4): 526-539.
- [20] Zhang L, Orloff M, Reber S, *et al*. cgaTOH: extended approach for identifying tracts of homozygosity[J]. *PLoS One*, 2013, 8(3): e57772.
- [21] Narasimhan V, Danecek P, Scally A, *et al*. BCFtools/ROH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data[J]. *Bioinformatics*, 2016, 32(11): 1749-1751.
- [22] Druet T, Gautier M. A model-based approach to characterize individual inbreeding at both global and local genomic scales[J]. *Mol Ecol*, 2017, 26(20): 5820-5841.
- [23] Saura M, Fernández A, Varona L, *et al*. Detecting inbreeding depression for reproductive traits in Iberian pigs using genome-wide data[J]. *Genet Sel Evol*, 2015, 47(1): 1.
- [24] Meyermans R, Gorssen W, Buys N, *et al*. How to study Runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species[J]. *BMC Genomics*, 2020, 21(1): 94.
- [25] Ai H S, Huang L S, Ren J. Genetic diversity, linkage disequilibrium and selection signatures in Chinese and Western pigs revealed by genome-wide SNP markers[J]. *PLoS One*, 2013, 8(2): e56001.
- [26] Zanella R, Peixoto J, Cardoso F, *et al*. Genetic diversity analysis of two commercial breeds of pigs using genomic and pedigree data[J]. *Genet Sel Evol*, 2016, 48(1): 24.
- [27] Zhang Z, Zhang Q, Xiao Q, *et al*. Distribution of Runs of homozygosity in Chinese and Western pig breeds evaluated by reduced-representation sequencing data[J]. *Anim Genet*, 2018, 49(6): 579-591.
- [28] Gorssen W, Meyermans R, Buys N, *et al*. SNP genotypes reveal breed substructure, selection signatures and highly inbred regions in Piétrain pigs[J]. *Anim Genet*, 2020, 51(1): 32-42.
- [29] Xu P, Wang X, Ni L, *et al*. Genome-wide genotyping uncovers genetic diversity, phylogeny, signatures of selection, and population structure of Chinese Jiangquhai pigs in a global perspective[J]. *J Anim Sci*, 2019, 97(4): 1491-1500.
- [30] Xie R, Shi L, Liu J, *et al*. Genome-wide scan for Runs of homozygosity identifies candidate genes in three pig breeds[J]. *Animals (Basel)*, 2019, 9(8): 518.
- [31] Huang M, Yang B, Chen H, *et al*. The fine-scale genetic structure and selection signals of Chinese indigenous pigs[J]. *Evol Appl*, 2019, 13(2): 458-475.
- [32] Schiavo G, Bovo S, Bertolini F, *et al*. Comparative evaluation of genomic inbreeding parameters in seven commercial and autochthonous pig breeds[J]. *Animal*, 2020, 14(5): 910-920.
- [33] Zhan H, Zhang S, Zhang K, *et al*. Genome-wide patterns of homozygosity and relevant characterizations on the population structure in Pietrain pigs[J]. *Genes (Basel)*, 2020, 11(5): 577.
- [34] Szmatała T, Jasielczuk I, Semik-Gurgul E, *et al*. Detection of Runs of homozygosity in conserved and commercial pig breeds in Poland[J]. *J Anim Breed Genet*, 2020, 137(6): 571-580.
- [35] Wang X, Zhang H, Huang M, *et al*. Whole-genome SNP markers reveal conservation status, signatures of selection, and introgression in Chinese Laiwu pigs[J]. *Evol Appl*, 2021, 14(2): 383-398.
- [36] Shi L, Wang L, Liu J, *et al*. Estimation of inbreeding and identification of regions under heavy selection based on Runs of homozygosity in a Large White pig population[J]. *J Anim Sci Biotechnol*, 2020, 11(1): 46.
- [37] Gorssen W, Meyermans R, Janssens S, *et al*. A publicly available repository of ROH islands reveals signatures of selection in different livestock and pet species[J]. *Genet Sel Evol*, 2021, 53(1): 2.
- [38] Kim E, Cole J, Huson H, *et al*. Effect of artificial selection on Runs of homozygosity in u.s. Holstein cattle[J]. *PLoS One*, 2013, 8(11): e80813.
- [39] 薛倩, 朱云芬, 李国辉, 等. 近交衰退分子机理的研究进展[J]. *家畜生态学报*, 2021, 42(4): 1-6.
- [40] Lopes M S, Silva F F, Harlizius B, *et al*. Improved estimation of inbreeding and kinship in pigs using optimized SNP panels[J]. *BMC Genet*, 2013, 14(1): 92.
- [41] Curik I, Ferenčaković M, Sölkner J. Inbreeding and Runs of homozygosity: A possible solution to an old problem[J]. *Livest Sci*, 2014, 166: 26-34.
- [42] Visscher P M, Medland S E, Ferreira M A, *et al*. Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings[J]. *PLoS Genet*, 2006, 2(3): e41.
- [43] McKay S D, Schnabel R D, Murdoch B M, *et al*. Whole genome linkage disequilibrium maps in cattle[J]. *BMC Genet*, 2007, 8: 74.
- [44] Leroy G. Inbreeding depression in livestock species: review and meta-analysis[J]. *Anim Genet*, 2014, 45(5): 618-628.
- [45] Marras G, Gaspa G, Sorbolini S, *et al*. Analysis of Runs of homozygosity and their relationship with inbreeding in five

- cattle breeds farmed in Italy[J]. *Anim Genet*, 2015, 46(2): 110-121.
- [46] Peripolli E, Stafuzza N, Munari D, *et al.* Assessment of Runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle[J]. *BMC Genomics*, 2018, 19(1): 34.
- [47] Giuffra E, Kijas J M, Amarger V, *et al.* The origin of the domestic pig: independent domestication and subsequent introgression[J]. *Genetics*, 2000, 154(4): 1785-1791.
- [48] Greger L, Keith D, Umberto A, *et al.* Worldwide phylogeography of wild boar reveals multiple centers of pig domestication[J]. *Science*, 2004, 307(5715): 1618-1621.
- [49] 国家畜禽遗传资源委员会. 中国畜禽遗传资源志: 猪志 [M]. 北京: 中国农业出版社, 2011: 1-10.
- [50] Martikainen K, Koivula M, Uimari P. Identification of Runs of homozygosity affecting female fertility and milk production traits in Finnish Ayrshire cattle[J]. *Sci Rep-Uk*, 2020, 10(1): 3804.
- [51] Falconer D, Mackay T. Inbreeding and crossbreeding: I. changes of mean value in introduction to quantitative genetics[M]. UK: Longman, 1996: 247-262.
- [52] Kardos M, Taylor H R, Ellegren H, *et al.* Genomics advances the study of inbreeding depression in the wild[J]. *Evol Appl*, 2016, 9(10): 1205-1218.
- [53] Martikainen K, Sironen A, Uimari P. Estimation of intrachromosomal inbreeding depression on female fertility using Runs of homozygosity in Finnish Ayrshire cattle[J]. *J Dairy Sci*, 2018, 101(12): 11097-11107.
- [54] Pryce J E, Haile-Mariam M, Goddard M E, *et al.* Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle[J]. *Genet Sel Evol*, 2014, 46(1): 71.
- [55] Kim E, Sonstegard T S, Van Tassell C P, *et al.* The relationship between Runs of homozygosity and inbreeding in Jersey cattle under selection[J]. *PLoS One*, 2015, 10(7): e129967.
- [56] Howard J, Tiezzi F, Huang Y, *et al.* A heuristic method to identify Runs of homozygosity associated with reduced performance in livestock[J]. *J Anim Sci*, 2017, 95(10): 4318-4332.

Current Knowledge and Application of Runs of Homozygosity in the Commercial Lean Pig Genome

LI Guixin, WANG Shiyuan, YANG Jie, WANG Xiaopeng*, ZHENG Enqin*

(College of Animal Science, South China Agricultural University, Guangdong Guangzhou 510642, China)

Abstract: Runs of homozygosity (ROH) are windows into population history and characteristic trait, and are widely used to predict wide-genome inbreeding levels and to identify selected genomic regions in livestock. Currently, the availability of genotypes for a large number of single nucleotide polymorphisms (SNPs) allows for more accurate characterisation of ROH, however, a reliable and uniform guideline for ROH analysis is lacking. In this study, we briefly describe the development history, detection software and parameters of ROH. Moreover, we provide a systematic review of the applications of ROH in the genetic breeding of commercial lean pigs, such as Duroc, Landrace, Yorkshire and Piétrain pigs which includes the assessment of genetic diversity, measurement of inbreeding coefficients, identification of artificial selection traces and detection of ROH genotypes associated with an unfavorable phenotype. Finally, we provide an insight into the development and application prospects of ROH.

Keywords: Lean pigs; ROH; Population history; Artificial selection; Unfavorable ROH

(责任编辑: 周会会)



中文核心期刊
中国科技核心期刊
RCCSE中国核心学术期刊
全国畜牧兽医优秀期刊

ISSN 1673-1182
CN 61-1433/S

家畜生态学报

ACTA ECOLOGIAE ANIMALIS DOMASTICI
JOURNAL OF DOMESTIC ANIMAL ECOLOGY

2019 第40卷 第5期
Vol.40 No.5



ISSN 1673-1182



西北农林科技大学主办

SPONSORED BY NORTHWEST A&F UNIVERSITY

※ 中文核心期刊(2011、2014、2017)
 ※ 中国科技核心期刊(中国科技论文统计源期刊)
 ※ RCCSE 中国核心学术期刊
 ※ 中国科学引文数据库源期刊
 ※ 《中国核心期刊(遴选)数据库》源期刊
 ※ 中国期刊全文数据库
 ※ 中国学术期刊综合评价数据库期刊
 ※ 中文科技期刊数据库源期刊
 ※ 《中国期刊网》、《中国学术期刊(光盘版)》固定源期刊
 ※ 《中国农业文摘-畜牧》固定引用源期刊
 ※ 《中文科技资料目录·农业》收录源期刊
 ※ 《CAJ-CD 规范》执行优秀奖
 ※ 全国畜牧兽医学类优秀期刊(第九届)
 ※ 陕西省优秀科技期刊一等奖
 ※ 陕西省优秀科技期刊(第四届)

主 编:陈玉林

副 主 编:杨雨鑫 陈小强

学 术 编 辑:张恩平 胡建宏 王 昕

杨雨鑫 郑惠玲 王永军

魏泽辉 杨朝霞 王小龙

曹阳春 宋宇轩 杨 欣

本期责任编辑:王 昕 陈小强

特约英文编辑:王丽萍

E-mail: jst@x263.net

网址: <http://jst.magtech.com.cn>

家畜生态学报

(月刊)

Journal of Domestic

Animal Ecology

(Monthly)

第 40 卷

Vol. 40

5

2019

主管单位:

中华人民共和国教育部

主办单位:

西北农林科技大学

协办单位:

中国畜牧兽医学会家畜生态学会

家畜生态学报

JIA CHU SHENG TAI XUE BAO

第40卷

第5期(总第204期)

2019年5月

学科动态

- 基因组水平的选择信号及其检测方法研究进展 王宇占,赵毅强(1)
- 日粮纤维水平和来源影响犊牛生长和胃肠道发育的研究 马满鹏,王炳,屠焰,等(7)

科学研究

- MIR-3880对奶山羊乳腺上皮细胞 OGR1 基因的靶向调控作用 侯金星,安小鹏,杜晓岩,等(13)
- 花生四烯酸影响绒山羊毛乳头细胞中维生素 D 受体基因表达的研究 高 晔,程幼敏,赵林虎,等(18)
- 奶牛溶菌酶基因(Lyz)乳腺特异性表达质粒构建与鉴定 张志鹏,江静宜,夏海磊,等(24)
- 低聚壳聚糖对北京鸭屠宰性能、血清抗氧化性能的影响 辛清武,李 丽,章琳俐,等(29)
- 益母草碱对肉仔鸡生长性能、免疫功能和肠道菌群的影响 乔彦杰,杨 莉,刘晓婷,等(34)
- 不同饲养模式对母猪分娩前后行为影响的研究 丁荣荣,张志伟,付帝生,等(39)
- 异食癖对哺乳期犊牛肠道黏膜通透性及粪便中大肠杆菌和双歧杆菌的影响 胡丹丹,金亚东,吉武龙,等(45)
- 日粮粗纤维水平对妊娠母猪规律行为及繁殖性能影响 张志伟,张 虎,黄大鹏(50)
- 不同 FSH 注射策略对澳洲白绵羊超数排卵效果的分析 卢小芳,苏布日,王 斌,等(55)
- 沼液施用量对大麦农艺性状及饲用品质的影响 杨智青,丁海荣,陈应江,等(59)
- 长期撂荒地植物种出现次数与个体数的空间异质性分析研究 冯兆佳,陈 俊,杨玉婷,等(66)

生态养殖

- 益生菌对八眉二元断奶仔猪抗腹泻情况的影响 马玉红,吴国芳(72)
- 牛舍有害气体排放规律及减除措施研究进展 刘 明,张恩平,宋宇轩(76)

专题论述

- 我国畜禽养殖防疫消毒设备研究应用现状 冯青春,张 俊,王 秀(82)
- 限位环境对妊娠母猪母性的影响 关立伟,于浪潮,赵晓宇,等(87)

教学园地

- “翻转课堂”教学模式在家畜寄生虫学实验教学中的应用 孙 莉,李 明(91)
- 基于 Mini-CEX 考核法的《兽医临床诊疗技术》教学改革 杨庆稳,张传师,雍 康,等(94)

[期刊基本参数] CN61-1433/S * 1080 * m * A1 * 95 * zh * P * ¥6.00 * 1000 * 49 * 2019-05

不同饲养模式对母猪分娩前后行为影响的研究

丁荣荣, 张志伟, 付帝生, 全建平, 王兴旺, 周身娉, 李绍云, 庄站伟,
蔡更元, 杨 杰, 吴珍芳, 刘德武, 郑恩琴*

(华南农业大学 动物科学学院, 国家生猪种业工程技术研究中心, 广东 广州 510642)

[摘 要] 为了研究限位栏饲养模式和群养模式对母猪在分娩前和分娩后行为的影响, 选取限位栏饲养模式下的大白和杜洛克怀孕母猪各 18 头、群养模式下的大白和杜洛克怀孕母猪各 10 头, 记录在分娩前 2 周以及分娩后 3 周内的母猪行为, 分析母猪分娩前后的行为差异以及不同饲养模式对母猪行为的影响。结果表明, 所有母猪行为中, 趴卧时间均占观察时间比例最高($>85\%$), 分娩后母猪趴卧时间增长($>93\%$), 而分娩前犬坐和站立时间占观察时间比例相对较少($<3.5\%$ 和 $<11.5\%$), 分娩后其占观察时间比例更小($<2.2\%$ 和 $<4.3\%$)。在两种饲养模式下, 两个品种母猪在分娩前的站立时间以及舔、嗅和针对饮水器等行为的频率均显著高于分娩后($P<0.05$), 分娩后的趴卧时间和针对饲槽的行为频率则显著更高($P<0.05$), 表明在两种饲养模式下, 母猪在分娩前与后的行为均表现出显著差异, 而杜洛克母猪分娩前与后的行为差异更明显。在分娩前, 限位栏饲养的两个品种母猪的啃栏和针对饮水器行为的频率显著高于群养母猪($P<0.05$), 而分娩后群养大白和杜洛克母猪的趴窝行为均显著高于限位栏饲养的大白和杜洛克母猪($P<0.05$), 表明限位栏饲养的母猪较群养母猪刻板行为发生频率更高, 福利水平较差。

[关键词] 怀孕母猪; 行为; 限位栏饲养; 群养

[中图分类号] S811.5

[文献标识码] A

[文章编号] 1005-5228(2019)05-0039-06

doi:10.3969/j.issn.1673-1182.2019.05.008

妊娠期母猪的饲养方式主要可分为大栏群养模式和限位栏饲养模式两种^[1]。限位栏饲养模式具有节约猪舍建筑面积、方便管理、价格便宜且便于观察母猪身体状况等优势^[2]; 群养模式则符合母猪群居生活的习性, 能够提供充分的活动空间和更加自由的行为表达而被生产管理者所重视。

2013 年开始欧盟禁止妊娠母猪使用限位栏进行饲养^[3-4], 动物福利逐渐被人们广泛关注。限位栏和群养模式都不可避免的带来空间限制问题^[1], 而空间限制会导致猪产生慢性应激甚至刻板行为^[5], 慢性应激和刻板行为都是猪生产中常见的动物福利问题^[6]。王国甫^[7]研究发现, 限位栏饲养对母猪的空嚼行为、非摄食口部行为、代谢行为均显著高于群养母猪, 而躺卧行为和社会行为则显著更低, 表明限位栏饲养明显增加了妊娠母猪的呆板行为, 减少其

社会行为的发生。顾宪红和张俊玲^[2]研究发现, 母猪群养模式中妊娠母猪的站立、躺卧、空口咀嚼、咬栏、排尿、饮水行为发生次数以及生殖器伤痕和眼鼻分泌物发生率显著下降, 睡觉、争斗、发声、采食和排便行为发生次数以及体表伤痕发生率显著升高, 表明了限位栏饲养模式应激更大, 福利更差。这与周勤^[8]和于韵青等^[9]研究结果相似。Rhodes 等^[10]发现, 群养条件下由于母猪间的相互争斗而损伤更容易引起母猪的恐惧和焦虑。且母猪被报道生产性能会出现下降^[11]。因此, 为确切的了解两种饲养模式对分娩前和分娩后母猪行为的影响和带来的动物福利问题, 本研究对分娩前后的母猪进行行为观察, 比较分娩前和分娩后母猪行为的差异, 分析不同饲养模式对母猪行为的影响, 以期改善妊娠期和哺乳期的母猪饲养管理和动物福利提供借鉴。

[收稿日期] 2018-02-27 **[修改日期]** 2018-06-26

[基金项目] 广东省自然科学基金(2017A030313213); 国家重点研发计划(2017YFD0502001)

[作者简介] 丁荣荣(1992—), 男, 江苏南通人, 博士研究生, 主要从事动物遗传育种研究; E-mail: 870281997@qq.com;

* **[通讯作者]** 郑恩琴(1981—), 女, 广东博罗人, 畜牧师, 硕士, 主要从事动物遗传育种研究。E-mail: eqzheng@scau.edu.cn

1 材料与方法

1.1 试验时间与地点

本试验于 2014 年~2016 年在华南农业大学和广东温氏食品集团股份有限公司联合组建的国家生猪种业工程技术研究中心完成。

1.2 试验材料

限位饲养母猪:怀孕经产母猪 36 头,怀孕日龄 95 d 左右。其中,大白母猪 18 头,杜洛克母猪 18 头;群养母猪:怀孕经产母猪 20 头,怀孕日龄 95 d 左右。其中,大白母猪 10 头,杜洛克母猪 10 头。

1.3 母猪的管理

1.3.1 限位栏母猪分娩前后的管理 怀孕舍限位栏规格为 $(0.15+1.95)\times 0.6\times 1.15\text{ m}^3$ 。怀孕母猪每头一栏,每天饲喂两次(上午 8:00,下午 16:30)。母猪产前 2~3 d 转入分娩舍产仔限位栏。怀孕舍采用喷雾降温。产仔限位栏为高床,分仔猪活动区、休息区和母猪趴卧区,仔猪活动区面积 $2\times(2.1\times 0.45)\text{ m}^2$,休息区 $1.15\times 0.6\text{ m}^2$,母猪趴卧区 $2.1\times 0.6\text{ m}^2$ 。栏内设有自动饮水器,仔猪补料槽,仔猪休息区上方设有红外线保温灯。哺乳母猪每天饲喂两

次(上午为 8:00,下午为 16:30)。分娩舍采用水帘降温负压通风系统。对分娩前后(分娩前 2 周)和哺乳期(分娩后 3 周)进行为期 5 周的行为观察。观察期间记录母猪的行为,尽量减少人为因素对母猪的影响。

1.3.2 群养母猪分娩前后的管理 母猪怀孕期间大栏群养,群养栏舍规格为 $4.8\times 3\text{ m}^2$,地面是水泥地面,有轻微坡度,排泄区和饮水区在下坡一侧,食槽在两侧,三面是围墙,一面是栏杆。每栏饲养 4 头怀孕母猪,每天饲喂两次(上午为 8:00,下午为 16:30)。采用水帘降温负压通风系统。母猪产前 2~3 d 将所有母猪转入至分娩舍产仔限位栏(大小、规格和管理方式同上)。

1.4 母猪的行为观察指标与判断标准

在预产期前 3 周开始进行行为观察直至母猪断奶结束,采用摄像头记录母猪行为,录像时间为每周—、三、五上午 9:00 到下午 16:00,并选择上午 9:00~10:30,中午 11:30~13:00,下午 14:00~15:30,3 个时间段视频进行行为分析。具体行为描述和记录方式见表 1。其中,X1、X2、X3 为状态行为,用时间计算;X4~X10 为事件行为,用频率计算。

表 1 母猪行为观察的相关行为参数

Table 1 Behavioral parameters related to sow behavior observation

行为代码 Behavior code	行为名称 Behavior	行为描述 Behavior description
X1	趴卧	包括睡眠趴卧和非睡眠趴卧
X2	站立	母猪四肢站立,包括走动站立和静止站立
X3	犬坐	前肢直立,后躯坐立在地板上
X4	嗅地	以鼻盘触地面嗅闻或掀拱包括站立嗅地、坐立嗅地和趴卧嗅地
X5	舔栏	母猪不断地舔舐铁栏
X6	针对饮水器的行为	母猪触碰饮水器及饮水的行为
X7	针对饲槽的行为	母猪接触饲槽以及吃料的行为
X8	啃栏	用牙齿啃咬栏内的围栏,包括站立啃栏、坐立啃栏和趴卧啃栏
X9	无食咀嚼	上下颚咬合,似乎在咀嚼食物,但是嘴里没有任何饲料和物体
X10	姿势转换的频率	观察时间内由一种姿势转换成另一种姿势的次数

1.5 统计分析

本试验所使用的统计分析软件为 R 语言。使用 anov 函数对不同饲养模式、不同品种母猪在分娩前和分娩后的母猪行为的表型数据进行组间方差分析。试验数据用平均数±标准误表示, $P<0.05$ 表示差异显著, $P>0.05$ 表示差异不显著。

2 结果与分析

2.1 分娩前后母猪的行为差异比较

由表 2 可知,在限位栏饲养模式下,大白和杜洛克母猪在分娩前的站立时间以及舔、嗅、啃栏、无食咀嚼和针对饮水器等行为的频率均显著高于分娩

后,分娩后的趴卧时间和针对饲槽的行为频率则显著更高($P<0.05$);两个品种不同之处则是,杜洛克母猪在分娩前的犬坐时间显著高于分娩后($P<0.05$),而大白母猪则差异不显著($P>0.05$)。在群养模式下,大白和杜洛克母猪在分娩前的站立时间和舔、嗅等行为的频率均显著高于杜洛克母猪分娩后的行为表现,而分娩后的趴卧时间和针对饲槽的行为频率则显著更高($P<0.05$),犬坐时间、啃栏、无食咀嚼和姿势转换则差异不显著($P>0.05$)。两品种的不同之处则是大白母猪在分娩前针对饮水器的行为频率显著高于分娩后($P<0.05$),而杜洛克则差异不显著($P>0.05$)。

表2 限位栏饲养和群养模式下大白和杜洛克母猪分娩前后的行为比较

Table 2 Comparison of behavior of the Yorkshire and Duroc sows before and after delivery under group housing and stall conditions

品种 Breed	母猪行为 Sows behavior	限位栏饲养 Stall housing		群养 Group housing	
		分娩前 Before delivery	分娩后 After delivery	分娩前 Before delivery	分娩后 After delivery
大白 Yorkshire	趴卧/%	88.79 ^a ±1.54	93.68 ^b ±0.75	87.66 ^a ±1.45	96.78 ^b ±0.51
	犬坐/%	3.13 ^a ±0.61	2.12 ^a ±0.42	2.51 ^a ±0.94	1.51 ^a ±0.37
	站立/%	7.68 ^b ±1.37	3.72 ^a ±0.60	9.94 ^b ±1.32	1.76 ^a ±0.32
	舔、嗅	1.05 ^b ±0.19	0.49 ^a ±0.09	1.37 ^b ±0.12	0.14 ^a ±0.04
	啃栏	0.45 ^b ±0.09	0.25 ^a ±0.07	0.11 ^a ±0.03	0.23 ^a ±0.05
	无食咀嚼	3.75 ^b ±0.22	2.50 ^a ±0.15	2.93 ^a ±0.28	2.32 ^a ±0.23
	针对饮水器	1.33 ^b ±0.18	0.44 ^a ±0.07	0.45 ^b ±0.07	0.16 ^a ±0.05
	针对饲槽	0.88 ^a ±0.13	2.06 ^b ±0.10	1.11 ^a ±0.14	1.51 ^b ±0.14
	姿势转换	3.71 ^a ±0.29	3.51 ^a ±0.41	3.40 ^a ±0.47	2.64 ^a ±0.40
杜洛克 Duroc	趴卧/%	85.64 ^a ±1.44	93.89 ^b ±0.93	90.76 ^a ±0.91	97.30 ^b ±0.61
	犬坐/%	3.34 ^b ±0.91	0.93 ^a ±0.28	0.81 ^a ±0.23	0.56 ^a ±0.22
	站立/%	11.31 ^b ±0.92	4.27 ^a ±0.58	8.37 ^b ±0.88	2.15 ^a ±0.51
	舔、嗅	1.73 ^b ±0.17	0.51 ^a ±0.10	1.16 ^b ±0.14	0.15 ^a ±0.03
	啃栏	0.66 ^b ±0.10	0.12 ^a ±0.04	0.13 ^a ±0.04	0.15 ^a ±0.03
	无食咀嚼	3.52 ^b ±0.22	1.93 ^a ±0.17	1.63 ^a ±0.12	1.78 ^a ±0.90
	针对饮水器	1.57 ^b ±0.14	0.43 ^a ±0.05	0.14 ^a ±0.04	0.26 ^a ±0.05
	针对饲槽	0.43 ^a ±0.08	1.55 ^b ±0.12	0.77 ^a ±0.09	1.66 ^b ±0.13
	姿势转换	3.53 ^a ±0.32	2.40 ^a ±0.28	2.04 ^a ±0.22	1.48 ^a ±0.30

注:同行肩标不同小写字母表示差异显著($P<0.05$)。下同。

Note: Values with different superscripts within a row mean significant difference ($P<0.05$). The same below.

2.2 不同饲养模式对母猪在分娩前和分娩后行为的影响

由表3可知,在分娩前,大白母猪和杜洛克母猪

在限位栏饲养模式下的啃栏和针对饮水器行为的频率均显著高于群养模式($P<0.05$);而杜洛克母猪在限位栏饲养模式下的趴卧和针对饲槽等行为的频

表3 分娩前和分娩后母猪在两种饲养模式下行为的比较

Table 3 Comparison of sows behavior before and after delivery under two feeding modes

品种 Breed	母猪行为 Sows behavior	分娩前 Before delivery		分娩后 After delivery	
		限位栏饲养 Stall housing	群养 Group housing	限位栏饲养 Stall housing	群养 Group housing
大白 Yorkshire	趴卧/%	88.79 ^a ±1.54	87.66 ^a ±1.45	93.68 ^a ±0.75	96.78 ^b ±0.51
	犬坐/%	3.13 ^a ±0.61	2.51 ^a ±0.94	2.12 ^a ±0.42	1.51 ^a ±0.37
	站立/%	7.68 ^a ±1.37	9.94 ^a ±1.32	3.72 ^b ±0.60	1.76 ^a ±0.32
	舔、嗅	1.05 ^a ±0.19	1.37 ^a ±0.12	0.49 ^b ±0.09	0.14 ^a ±0.04
	啃栏	0.45 ^b ±0.09	0.11 ^a ±0.03	0.25 ^a ±0.07	0.23 ^a ±0.05
	无食咀嚼	3.75 ^a ±0.22	2.93 ^a ±0.28	2.50 ^a ±0.15	2.32 ^a ±0.23
	针对饮水器	1.33 ^b ±0.18	0.45 ^a ±0.07	0.44 ^b ±0.07	0.16 ^a ±0.05
	针对饲槽	0.88 ^a ±0.13	1.11 ^a ±0.14	2.06 ^b ±0.10	1.51 ^a ±0.14
	姿势转换	3.71 ^a ±0.29	3.40 ^a ±0.47	3.51 ^a ±0.41	2.64 ^a ±0.40
杜洛克 Duroc	趴卧/%	85.64 ^a ±1.44	90.76 ^b ±0.91	93.89 ^a ±0.93	97.30 ^b ±0.61
	犬坐/%	3.34 ^b ±0.91	0.81 ^a ±0.23	0.93 ^b ±0.28	0.56 ^a ±0.22
	站立/%	11.31 ^a ±0.92	8.37 ^a ±0.88	4.27 ^b ±0.58	2.15 ^a ±0.51
	舔、嗅	1.73 ^a ±0.17	1.16 ^a ±0.14	0.51 ^b ±0.10	0.15 ^a ±0.03
	啃栏	0.66 ^b ±0.10	0.13 ^a ±0.04	0.12 ^a ±0.04	0.15 ^a ±0.03
	无食咀嚼	3.52 ^b ±0.22	1.63 ^a ±0.12	1.93 ^b ±0.17	1.78 ^a ±0.90
	针对饮水器	1.57 ^b ±0.14	0.14 ^a ±0.04	0.43 ^a ±0.05	0.26 ^a ±0.05
	针对饲槽	0.43 ^a ±0.08	0.77 ^b ±0.09	1.55 ^a ±0.12	1.66 ^a ±0.13
	姿势转换	3.53 ^b ±0.32	2.04 ^a ±0.22	2.40 ^a ±0.28	1.48 ^a ±0.30

率显著低于群养模式,限位栏饲养模式下的犬坐、无食咀嚼和姿势转换等行为的频率则显著高于群养模式($P<0.05$),大白母猪则没有表现出明显的显著

性差异($P>0.05$)。在分娩后,大白母猪在限位栏饲养模式下的站立、舔、嗅、针对饮水器和针对饲槽等行为的时间和频率显著高于群养母猪($P<$

0.05), 群养大白母猪的趴窝行为则显著高于限位栏饲养的大白母猪 ($P < 0.05$); 而杜洛克母猪在限位栏饲养模式下的犬坐、站立、舔、嗅、无食咀嚼和姿势转换等行为显著高于群养杜洛克母猪, 趴卧行为的时间比例则显著低于群养杜洛克母猪 ($P < 0.05$)。

3 讨 论

3.1 分娩前、后母猪行为表现的一般规律

不同生长时期, 猪会通过调整或者改变行为来适应周围环境的变化, 如在分娩后, 母猪会通过增加更多趴卧行为来满足仔猪的营养需求。正常情况下, 母猪绝大部分时间处于趴卧状态^[12]。本研究中, 无论大白母猪还是杜洛克母猪, 分娩前趴卧时间占据观察总时间 87% 以上, 分娩后达到 93% 以上, 均占据花费时间的主要部分, 而犬坐和站立时间占观察时间比例相对较少, 这和崔卫国^[13], 刘洪贵^[14], 尹国安^[15]和 Chapinal 等^[16]报道的趴卧行为占观察时间的主要部分的结果类似。母猪在分娩后表现出更多趴卧行为和更少的站立和犬坐行为, 这主要与母猪在分娩后花费更多时间来给仔猪授乳相关。

随着猪场集约化生产的快速发展, 饲养密度高度增加, 猪场空间被高度利用, 不可避免的会对猪造成空间限制, 而空间限制会对猪的生理和心理造成慢性应激, 导致规癖行为^[9,16]。规癖行为指动物长时间地、有规则的重复某种没有明显生物学功能的固定模式的行为^[17]。在空间受限等限制条件下(限位栏), 母猪会表现出无食咀嚼、啃栏和针对饮水器等规癖行为^[18-19]。在本研究中, 限位栏饲养和群养两种饲养模式下母猪均出现了不同程度规癖行为, 这可能是由于群养模式下试验的环境丰富度不够导致。本研究发现, 除针对饲槽外, 啃栏、无食咀嚼和针对饮水器等其他规癖行为的频率在分娩后均有下降趋势。妊娠期是一个长期的过程, 这导致妊娠母猪长期处于压抑的状态, 并在产前 3 周达到最高峰, 母猪长期处于紧张和焦虑的状态会导致规癖行为发生的频率增加^[20]。而母猪在分娩后, 由于仔猪的存在, 一定程度上缓解了母猪压抑的状态, 规癖行为发生的频率降低。对于母猪在分娩后针对饲槽的行为频率显著增加, 分娩后的母猪需要更多的采食来维持哺乳及身体恢复所需要的营养, 这导致母猪针对饲槽的行为的频率和时间增多。不同品种的母猪行为之间也存在差异, 刘洪贵^[14]和 Valros 等^[21]也研究表明不同品种的母猪在哺乳期的行为存在差异。邓红雨^[22]研究发现, 不同品种母猪在哺乳阶段的趴

卧时间、趴卧位置和趴卧姿势均存在显著差异。本研究发现, 杜洛克相较于大白母猪在分娩前后的行为差异更明显, 这可能是由于品种的特异性导致的, 在生产中需要特别关注这一现象。

3.2 不同饲养模式对分娩前后母猪行为的影响

限位栏饲养因节约猪舍建筑面积和方便管理等优势已成为妊娠母猪主要的饲养模式^[23-24], 但限位栏限制了母猪的运动技能且对母猪的动物福利方面产生不利影响^[25]。本研究发现, 无论是大白母猪还是杜洛克母猪, 在分娩前和分娩后相较于群养模式, 限位栏饲养的母猪均表现出更频繁的规癖行为, 这与周勤^[26]和张俊玲等^[1]的研究结果一致。群养模式符合母猪群居动物的习性, 而限位栏抑制了母猪的活动, 母猪的正常行为得不到表达, 在这种情况下, 母猪更容易产生各种规癖行为^[20], 只能借助无食咀嚼、针对饮水器等规癖行为来释放这种限制带来的压力和不安^[27]。

限位栏饲养模式和群养模式饲养的母猪在预产期前 3 d 会转入至分娩舍产仔限位栏。这些母猪接受统一的管理与饲养, 理论上母猪行为上差异并不大, 然而, 分娩后限位栏饲养小组和群养小组的母猪行为依旧存在显著差异, 如无食咀嚼、针对饮水器和针对饲槽等规癖行为。在上述规癖行为中, 无食咀嚼可能是最有害的规癖行为, 可以作为环境束缚的一个重要参考指标^[8]。有研究表明, 通过改变环境很难消除甚至减少长期建立的规癖行为如无食咀嚼^[28]。这也暗示长期的限位栏饲养模式所带来的母猪规癖行为会传递到其他饲养模式中去, 甚至影响母猪终身的行为表现。

站立、趴卧等姿势行为被认为主要与动物的舒适度相关^[1]。不少研究均表明, 群养模式下母猪站立行为发生频率显著低于限位栏饲养模式^[29-30], 这与本研究发现的分娩后杜洛克和大白母猪的表现结果一致。Chapinal 等^[16]提出, 站立行为是母猪处于警觉状态的一种表现形式。本研究还发现, 分娩前后杜洛克母猪在群养模式下的趴卧行为显著高于限位栏饲养模式。许多研究发现, 在受限制环境生活的母猪会表现出更多的犬坐和站立的行为, 而趴卧行为表现较少, 所以趴卧行为被认为是母猪舒适的一种表现^[31-32]。然而, 这种结论也存在争议, 本研究还发现, 在母猪趴卧的同时往往伴随着无食咀嚼现象。这可能正是环境和空间的刺激导致母猪降低了姿势变换的频率, 而是长时间的保持一种姿势如长时间的趴卧和无食咀嚼, 这也可能是母猪适应不利

环境的一种表现。Anil 等^[33]也提出相似的观点,认为一些姿势行为指标如站立、趴卧等并不能用于表示圈舍系统内怀孕母猪的舒适情况,因为这些指标被认为是母猪适应环境束缚的策略。此外,在杜洛克母猪和大白母猪中,限位栏饲养的母猪的姿势转换频率均高于群养模式,这可能与限位栏饲养母猪为了缓解应激,通过不断调整姿势来提高身体的舒适度相关。尽管姿势行为指标能否作为监测母猪舒适情况的参考依据还有待进一步研究,可能还需要综合其他各方面的因素来评估母猪舒适和福利情况,如母猪的激素水平作为评估指标。将皮质醇作为一种应激指标已逐渐被人所接受,Estienne 等^[34]研究发现,限位栏饲养的母猪在妊娠 30 d 时的血清皮质醇的浓度要高于群养模式。于韵青等^[9]研究发现限位栏母猪的唾液淀粉酶与无食咀嚼行为之间存在显著的正相关关系。这提示在后期评估母猪慢性应激状态或者心理状态可增加激素水平的检测作为参考。就目前的研究来看,限位栏饲养母猪福利低于群养母猪,群养模式更有利于母猪行为的充分表达。

4 结 论

限位栏饲养和群养两种模式下,母猪在分娩前和分娩后的行为均表现出显著的差异。就品种而言,杜洛克母猪分娩前和分娩后的行为表现出的差异更明显。限位栏饲养的母猪较群养母猪规癖行为发生频率更高,福利水平较差,且长期的限位栏饲养模式所带来的母猪规癖行为表现出能够传递到其他饲养模式中去的趋势。

参考文献:

- [1] 张俊玲,顾宪红. 妊娠母猪限位栏和群饲系统比较[J]. 家畜生态学报, 2016, 37(2): 81-85.
- [2] 顾宪红,张俊玲. 母猪电子群养系统和个体限位栏系统繁殖性能及动物福利水平对比分析[J]. 畜牧兽医学报, 2016, 47(6): 1 189-1 197.
- [3] 叶娜,黄川. 荷兰 Velos 智能化母猪饲养管理系统在国内猪场的应用[J]. 养猪, 2009(2): 41-42.
- [4] 黄河龙. 2013 年部分欧盟成员国暂不能禁用母猪限位栏[J]. 猪业科学, 2012(7): 24.
- [5] 王昕陟,包军. 妊娠母猪的规癖行为与采食动机的关系(综述)[J]. 养猪, 2003(1): 10-12.
- [6] Spoolder H A M, Geudeke M J, Van der Peet-Schwering C M C, et al. Group housing of sows in early pregnancy: A review of success and risk factors[J]. Livestock Science, 2009, 125(1): 1-14.
- [7] 王国甫. 母体妊娠期活动受限对母体及子代行为和代谢指标的影响[D]. 南京: 南京农业大学, 2013.
- [8] 周勤. 群养与限位栏饲养模式下妊娠母猪繁殖性能及行为的比较研究[D]. 南京: 南京农业大学, 2012.
- [9] 于韵青,包军. 不同饲养模式对妊娠母猪行为、唾液皮质醇和淀粉酶水平的影响[J]. 中国畜牧杂志, 2013, 49(13): 69-73.
- [10] RHODES R T, APPLEBY M C, CHINN K, et al. Task Force Report -A comprehensive review of housing for pregnant sows[J]. Javma-Journal of the American Veterinary Medical Association, 2005, 227(10): 1 580-1 590.
- [11] BARNETT J L, HEMSWORTH P H, CRONIN G M, et al. A review of the welfare issues for sows and piglets in relation to housing[J]. Australian Journal of Agricultural Research, 2001, 52(1): 1-28.
- [12] LOU Z, HURNIK J F. Peripartum sows in three farrowing crates: Posture patterns and behavioural activities[J]. Applied Animal Behaviour Science, 1998, 58(1-2): 77-86.
- [13] 崔卫国. 不同生产环境和遗传因素对母猪行为规癖影响的研究[D]. 哈尔滨: 东北农业大学, 2002.
- [14] 刘洪贵. 不同福利措施及品种对母猪的行为、生理、免疫、健康及生产性能的影响[D]. 哈尔滨: 东北农业大学, 2013.
- [15] 尹国安. 不同畜舍环境对猪的生产性能、行为表达及生理状况的影响[D]. 哈尔滨: 东北农业大学, 2010.
- [16] CHAPINAL N, RUIZ DE LA TORRE J L, CERISUELO A, et al. Evaluation of welfare and productivity in pregnant sows kept in stalls or in 2 different group housing systems [J]. Journal of Veterinary Behavior-Clinical Applications and Research, 2010, 5(2): 82-93.
- [17] NOWICKI J, KLOCEK C, SCHWARZ T. Factors affecting maternal behaviour and responsiveness in sows during periparturient and lactation periods [J]. Annals of Animal Science, 2012, 12(4): 455-469.
- [18] DANTZER R. Behavioral, physiological and functional aspects of stereotyped behavior: a review and a re-interpretation[J]. Journal of Animal Science, 1986, 62(6): 1 776-1 786.
- [19] 王宏玲,廖新佛,王成跃,等. 土种母猪和洋母猪的行为比较研究[J]. 家畜生态学报, 2006, 27(05): 57-59.
- [20] 周鹏. 限位栏饲养对母猪的损害[J]. 养殖技术顾问, 2011(4): 60-61.
- [21] VALROS A, RUNDGREN M, SPINKA M, et al. Metabolic state of the sow, nursing behaviour and milk production[J]. Livestock Production Science, 2003, 79: 155-167.
- [22] 邓红雨. 产仔限位栏和分娩圈中母猪分娩前后行为的研究[D]. 哈尔滨: 东北农业大学, 2001.
- [23] 李永辉. 智能化母猪群养在中国养猪生产中的应用[J]. 中国猪业, 2011(9): 4-7.
- [24] 涂金敏,傅金奎,叶健,等. 电子母猪群养系统和限位栏系统猪繁殖性能的比较[C]//福建厦门: 中国猪业科技大会暨中国畜牧兽医学学会 2015 年学术年会, 2015.
- [25] 丁立军. 智能母猪群养系统在北京绿都种猪育种有限公司的应用[J]. 猪业科学, 2013, (10): 82-84.

- [26] 周勤, 王国甫, 周波, 等. 群养与限位栏饲养模式下妊娠母猪繁殖性能及行为的比较研究[C]//辽宁沈阳:全国动物生理生化第七届全国代表大会暨第十三次学术交流会论文摘要汇编, 2014.
- [27] CRONIN G M, SIMPSON G J, HEMSWORTH P H. The effects of the gestation and farrowing environments on sow and piglet behaviour and piglet survival and growth in early lactation[J]. *Applied Animal Behaviour Science*, 1996, 46(3-4): 175-192.
- [28] MASON G J. Stereotypies and suffering[J]. *Behavioural Processes*, 1991, 25(2-3): 103-115.
- [29] DEN HARTOG L A, BACKUS G B, VERMEER H M. Evaluation of housing systems for sows[J]. *Journal of Animal Science*, 1993, 71(5): 1 339-1 344.
- [30] HARRIS M J, PAJOR E A, SORRELLS A D, et al. Effects of stall or small group gestation housing on the production, health and behaviour of gilts[J]. *Livestock Science*, 2006, 102(1-2): 171-179.
- [31] HALES J, MOUSTSEN V A, NIELSEN M B F, et al. The effect of temporary confinement of hyperprolific sows in Sow Welfare and Piglet protection pens on sow behaviour and salivary cortisol concentrations[J]. *Applied Animal Behaviour Science*, 2016, 183: 19-27.
- [32] MORRIS J R, HURNIK J F, FRIENDSHIP R M, et al. The behavior of gestating swine housed in the Hurnik-Morris system[J]. *J Anim Sci*, 1993, 71(12): 3 280-3 284.
- [33] ANIL L, ANIL S S, DEEN J. Relationship between postural behaviour and gestation stall dimensions in relation to sow size[J]. *Applied Animal Behaviour Science*, 2002, 77(3): 173-181.
- [34] ESTIENNE M J, HARPER A E, KNIGHT J W. Reproductive traits in gilts housed individually or in groups during the first thirty days of gestation[J]. *Journal of Swine Health and Production*, 2006, 14(5): 241-246.

Effects of Different Feeding Patterns on Sows' Behavior Before and After Delivery

DING Rongrong, ZHANG Zhiwei, FU Disheng, QUAN Jianping, WANG Xingwang, ZHOU Shenping, LI Shaoyun, ZHUANG Zhanwei, CAI Gengyuan, YANG Jie, WU Zhenfang, LIU Dewu, ZHENG Enqin*

(National Engineering Research Center for Breeding Swine Industry, College of Animal Science, South China Agricultural University, Guangzhou 510642)

Abstract: The aim of this study was to analyze the effects of group housing and stall condition modes on behavior of sows before and after delivery. In the study, 36 pregnant sows were selected under the stall conditions (18 Yorkshires and 18 Durocs), 20 pregnant sows were selected under the group housing (10 Yorkshires and 10 Durocs), and using the camera to record them for 2 weeks before delivery and 3 weeks after delivery. The behavioral differences of sows before and after delivery and the effects of different feeding patterns on sows behavior were analyzed. In all the behaviors, the duration of lying was the longest in the observation time ($>85\%$), and the sow's lying duration increased after delivery ($>93\%$). However, the ratio of sitting and standing time before delivery was relatively small ($<3.5\%$ and $<11.5\%$), and the proportion of observation time was smaller after delivery ($<2.2\%$ and $<4.3\%$). In two different feeding modes, standing time and the frequency of licking, sniffing and drinking fountains of all sows before delivery was significantly higher than that after delivery ($P<0.05$), and postpartum lying time and manger frequency were significantly higher ($P<0.05$), suggesting a significant difference in the behavior of sows before and after delivery in both group housing and stall conditions. Prior to delivery, the behavior of biting bars and aiming at drinker under the stall conditions was significantly higher than that of the group housing ($P<0.05$). After delivery, lying behavior of the sows raised in the group housing, which was significantly higher than that in the stall conditions ($P<0.05$), indicating that the sows raised in the stall conditions had higher frequency of stereotypic behavior than group housing sows, with poor welfare.

Key words: pregnant sow; behavior; stall conditions; group housing

- ◎ 中文核心期刊
- ◎ 中国科技核心期刊(中国科技论文统计源期刊)
- ◎ RCCSE中国核心学术期刊
- ◎ 《中国核心期刊(遴选)数据库》源期刊
- ◎ 中国科学引文数据库源期刊
- ◎ 中国期刊全文数据库
- ◎ 中国学术期刊综合评价数据库源期刊
- ◎ 中文科技期刊数据库源期刊
- ◎ 《中国农业文摘—畜牧》固定引用源期刊
- ◎ 《中国期刊网》、《中国学术期刊(光盘版)》固定源期刊

欢迎投稿 欢迎订阅

2019年《家畜生态学报》

《家畜生态学报》是由中华人民共和国教育部主管,西北农林科技大学主办的面向国内外公开发行的科技期刊,主要刊登家畜生态、家畜环境控制及牧业生态经济研究方面取得的创新成果,介绍畜牧生态学科国内外研究与发展动态,报道畜禽资源评价、保护与利用的理论与实践成果,以促进家畜生态学科基础理论和应用技术的创新发展,服务于国家生态与经济建设。本刊为月刊,每期内文为96页码,A4开本,每期定价6.00元,全年订价72.00元,本刊代号为52-112,当地邮局可办理订购手续,亦可直接向《家畜生态学报》编辑部订阅。

地址:陕西杨凌西北农林科技大学动物科技学院《家畜生态学报》编辑部;邮编:712100;

电话(传真):(029)87091130;电子信箱:jcst@x263.net;联系人:陈小强

网址: <http://jcst.magtech.com.cn>

家畜生态学报

JIA CHU SHENG TAI XUE BAO

月刊 1980年创刊

第40卷 第5期 2019年5月

ACTA ECOLOGIAE ANIMALIS DOMASTICI

Monthly, Started in 1980

Vol.40 No.5 May 2019

主管单位: 中华人民共和国教育部

主办单位: 西北农林科技大学

主 编: 陈玉林

编辑出版:《家畜生态学报》编辑部

(陕西杨凌西北农林科技大学

动物科技学院,邮编712100,

电话(传真): 029-87091130)

印刷单位: 陕西森奥印务有限公司

国内发行: 全国各地邮局(所)

(本刊代号52-112)

Authorized by the Ministry of Education of the
People's Republic of China

Sponsored by Northwest A & F University

Chief Editor CHEN Yu-lin

Edited & Published by Editorial Department of

Journal of Domestic Animal Ecology (College

of Animal Science and Technology, North-

west A & F University, Yangling, Shaanxi,

712100, China). Tel(Fax): (029)87091130

Printed by Senao Printing Co., Ltd. Shaanxi

Distributed within China by Post Offices through-
out the Country

ISSN 1673-1182

CN 61-1433/S

本期责任编辑: 王 昕 陈小强



本期英文编辑: 王丽萍

每期定价: 6.00元

全年定价: 72.00元

Article

Gender Control of Mouse Embryos by Activation of TLR7/8 on X Sperm via Ligands dsRNA-40 and dsRNA-DR

Yunfei Hou ^{1,2,3,4,5} , Jingfeng Peng ^{1,2,3,4,5}, Linjun Hong ^{1,2,3,4,5} , Zhenfang Wu ^{1,2,3,4,5,6}, Enqin Zheng ^{1,2,3,4,5,*} and Zicong Li ^{1,2,3,4,5,6,*}

- ¹ National Engineering Research Center for Breeding Swine Industry, South China Agricultural University, Guangzhou 510642, China; hyf2021@stu.scau.edu.cn (Y.H.); 20231024011@stu.scau.edu.cn (J.P.); linjun.hong@scau.edu.cn (L.H.); wzf@scau.edu.cn (Z.W.)
 - ² State Key Laboratory of Swine and Poultry Breeding Industry, South China Agricultural University, Guangzhou 510642, China
 - ³ National and Local Joint Engineering Research Center for Livestock and Poultry Breeding Industry, South China Agricultural University, Guangzhou 510642, China
 - ⁴ Department of Animal Genetics, Breeding and Reproduction, College of Animal Science, South China Agricultural University, Guangzhou 510642, China
 - ⁵ Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, South China Agricultural University, Guangzhou 510642, China
 - ⁶ Guangdong Provincial Laboratory of Lingnan Modern Agricultural Science and Technology, Guangzhou 510642, China
- * Correspondence: eqzheng@scau.edu.cn (E.Z.); lizicong@scau.edu.cn (Z.L.)

Abstract: Gender control technologies are promising for enhancing the production efficiency of the farm animal industry, and preventing sex-linked hereditary diseases in humans. It has been shown that the X sperm of mammalian animals specifically expresses X-chromosome-derived toll-like receptor 7/8 (TLR7/8), and the activation of TLR7/8 on the X sperm by their agonist, R848, can separate X and Y sperm via the specific inhibition of X sperm motility. The use of R848-preselected sperm for fertilization resulted in sex-ratio-skewed embryos or offspring. In this study, we aimed to investigate whether two other TLR7/8 ligands, double-stranded RNA-40 (dsRNA-40) and double-stranded RNA-DR (dsRNA-DR), are also effective in the separation of mouse X and Y sperm and the subsequent generation of gender-ratio-skewed in vitro fertilization (IVF) embryos. Our results indicated that cholesterol modification significantly enhances the transfection of dsRNA-40 and dsRNA-DR into sperm cells. dsRNA-40 and dsRNA-DR incubation with mouse sperm could separate X and Y sperm by the specific suppression of X sperm motility by decreasing its ATP level and mitochondrial activity. The use of a dsRNA-40- or dsRNA-DR-preselected upper layer of sperm, which predominantly contains high-motility Y sperm, for IVF caused a male-biased sex ratio shift in resulting embryos (with 65.90–74.93% of embryos being male). This study develops a simple new method for the efficient separation of mammalian X and Y sperm, enabling the selective production of male or female progenies.

Keywords: gender control; sex ratio; dsRNA; TLR7/8; sperm separation



Citation: Hou, Y.; Peng, J.; Hong, L.; Wu, Z.; Zheng, E.; Li, Z. Gender Control of Mouse Embryos by Activation of TLR7/8 on X Sperm via Ligands dsRNA-40 and dsRNA-DR. *Molecules* **2024**, *29*, 262. <https://doi.org/10.3390/molecules29010262>

Received: 5 December 2023

Revised: 27 December 2023

Accepted: 29 December 2023

Published: 4 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In animal husbandry, the gender ratio is an important factor that can affect production efficiency. For example, in the dairy and poultry egg production industry, a higher female-to-male sex ratio is favorable for boosting milk yield and egg production [1,2]. Selective production of male or female breeding stock is beneficial for optimizing the reproductive potential of the parent population [3]. In humans, some hereditary diseases are transmitted via a gender-associated pattern [4,5]. Therefore, developing technologies for preselecting offspring sex has significant practical value in enhancing the production efficiency of the farm animal industry and reducing the transmission of sex-linked genetic diseases in humans.

At present, mammalian animal gender control can be achieved by flow-cytometry-based X/Y sperm separation and gene editing. Flow cytometry distinguishes X and Y sperm according to sex chromosome DNA content [6]. This method can preselect progeny sex with an accuracy of approximately 90% in farm animals like cattle, horses, sheep, and pigs [7–12]. However, it requires specialized equipment and expertise, with costs directly proportional to the number of separated sperm. This limitation makes it unsuitable for animals requiring a large dose of sperm for artificial insemination, such as pigs [13]. CRISPR/Cas9-mediated gene editing is a new, emerging technique for sex control. In 2019, Yosef et al. [14] demonstrated a CRISPR/Cas9 system for biasing mouse offspring ratios via the inactivation of developmental vital genes, specifically in male embryos. Douglas et al. [15] employed a synthetic CRISPR-Cas9 approach to producing a single-sex litter, albeit with a 30–40% reduction in litter size, by the specific elimination of male or female mouse embryos. In 2022, Bai [16] expressed a CRISPR/Cas9 system during mouse spermatogenesis to target the Y chromosome in Y sperm, generating female-biased progenies. Although these gene-editing-based technologies are promising for gender control, their commercial viability is limited by ethical and biosafety concerns related to genetic modifications. Hence, new techniques for animal gender control need to be developed.

In 2019, Umehara et al. [17] conducted a groundbreaking study, outlining a new method for the in vitro separation of X and Y sperm in mice using a TLR7/8 agonist, R848. Their research demonstrated that the activation of X-sperm-specific receptors TLR7/8 by the ligand R848 affected the glycolytic pathway and citric acid cycle, leading to reduced ATP levels and subsequent motility suppression. Y sperm, lacking TLR7/8, remained unaffected in terms of motility. This difference in X/Y sperm motility enabled their separation. Umehara later applied this system to separate frozen cattle semen, with similar results [18], offering a fresh perspective on animal sex control technology. In 2021, Chinese scientists led by Ren [19] applied this approach to treat goat sperm with R848, resulting in successful artificial insemination that produced sexually oriented young. Their research elucidated that R848 treatment inhibited X sperm motility in goats through the TLR7/8 signaling pathway, affecting X sperm ATP levels via the GSK3 α / β -hexokinase pathway. In 2022, Huang [20] combined R848 with an alkaline semen diluent, achieving the successful isolation and enrichment of goat X sperm. This sex control technique, relying on the R848 activation of X sperm TLR7/8, was successfully validated in several mammals, including mice, cattle, and sheep, confirming the robustness of this sex control technology [17–21].

R848 is highly stable and can induce a long-term immune response [22]. It may cause unwanted side effects when used to treat sperm. In addition to R848, double-stranded (ds) RNA-40 and dsRNA-DR also serve as TLR7/8 ligands [23–27], and both dsRNA-40 and dsRNA-DR have the potential to activate TLR7 and TLR8 (TLR7/8). However, whether these ligands can specifically inhibit X sperm motility and can be employed for sex control are unknown. In this study, we investigated the effects of dsRNA-40 and dsRNA-DR on mouse sperm motility, mitochondrial membrane potential, ATP levels, the proportion of X/Y sperm in the upper semen layer, and the sex ratio of embryos generated by in vitro fertilization with dsRNA-treated semen. The findings from this study contribute to the development of novel methods for animal sperm separation and sex control.

2. Results

2.1. Efficient Transfection of Sperm with Cholesterol-Modified dsRNA-40 and dsRNA-DR

TLR7/8 is usually located in the endosome inside eukaryotic cells [28–30]. Cholesterol modification has been shown to enhance the transfection of dsRNA into eukaryotic cells [31–33]. To facilitate the smooth entry of dsRNA into sperm to bind to TLR7/8, we used cholesterol-modified dsRNAs labeled with FAM fluorophore to incubate with mouse sperm (see Figure 1). Under fluorescence microscopy, green fluorescence was observed in sperm transfected with cholesterol-modified dsRNA-40 and dsRNA-DR, but not in sperm incubated with dsRNAs lacking cholesterol modification. These results suggest that cholesterol-modified dsRNA can efficiently penetrate sperm.

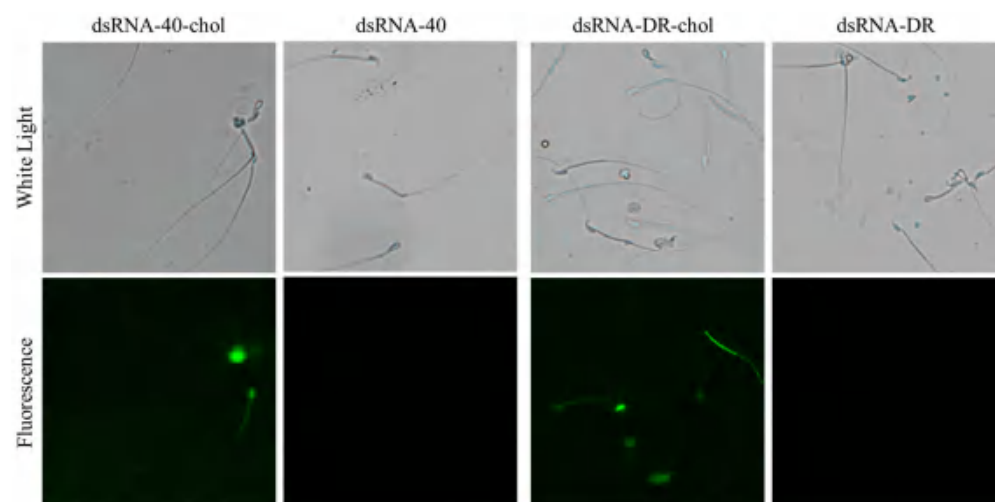


Figure 1. Efficient transfection of sperm with cholesterol-modified dsRNA-40 and dsRNA-DR.

2.2. Determination of Incubation Concentration and Time of dsRNA-40 and dsRNA-DR for Treatment of Mouse Sperm

To investigate the impact of TLR7/8 agonists dsRNA-40 and dsRNA-DR on mouse sperm motility, various concentrations of dsRNA-40 or dsRNA-DR were incubated with sperm for 1 h, followed by the assessment of sperm motility. The results presented in Figure 2a demonstrate that treatment with 0.3, 1, 3 and 10 μM dsRNA-40 or dsRNA-DR led to a significant reduction in sperm motility, and this effect exhibited concentration dependence. These data suggest that if employing dsRNA-40 and dsRNA-DR as TLR7/8 ligands to specifically inhibit the motility of X sperm, a concentration of 0.3 μM or higher should be used. Taking into account the findings from Umehara's study [17] that using 0.3 μM of R848, another TLR7/8 agonist, to incubate with mouse sperm is effective in separating X and Y sperm, we decided to utilize 0.3 μM dsRNA for the subsequent treatment of mouse sperm.

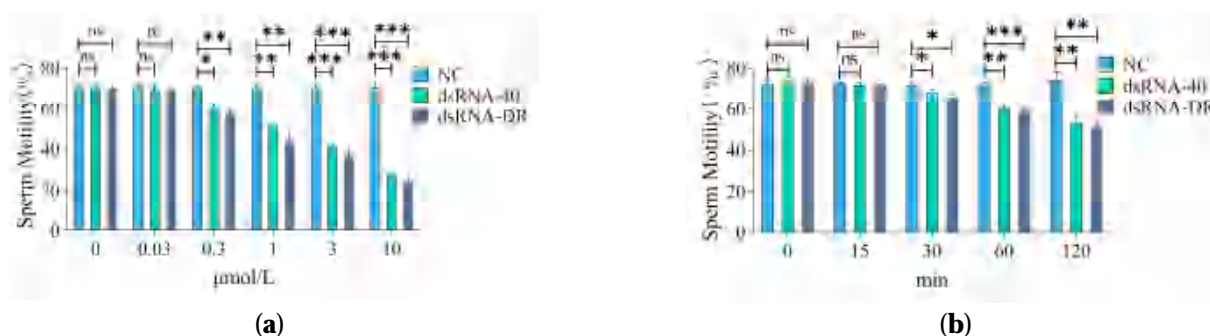


Figure 2. Determination of incubation concentration and time of dsRNA-40 and dsRNA-DR for treatment of mouse sperm. (a) Sperm motility was assessed following a 60 min incubation with varying concentrations of dsRNA-40 or dsRNA-DR; (b) sperm motility after incubation with 0.3 μM dsRNA-40 or dsRNA-DR for different times. Values represent the mean of three replicates \pm Standard Error of the Mean (SEM). NC denotes non-sense double-stranded RNA as control. ns denotes no statistically significant difference compared to control group, $p > 0.05$; * denotes a significant difference compared to the control group, $p < 0.05$; ** denotes a highly significant difference compared to the control group, $p < 0.01$; *** denotes an extremely significant difference compared to the control group, $p < 0.001$.

To further investigate the time required for 0.3 μM dsRNA treatment to alter sperm motility, mouse sperm were incubated with 0.3 μM dsRNA, and sperm motility was assessed at various timepoints. The results presented in Figure 2b demonstrate that treatment with 0.3 μM dsRNA-40 or dsRNA-DR for 30, 60, and 120 min significantly decreased sperm

motility, and this effect displayed a time-dependent pattern. These findings imply that the treatment of mouse sperm with 0.3 μ M dsRNA-40 or dsRNA-DR for 30 min may be effective in reducing X sperm motility. To avoid the detrimental effects caused by a longer treatment with dsRNA on sperm, we chose 30 min as the incubation time for subsequent dsRNA treatment of mouse sperm.

2.3. dsRNA-40 and dsRNA-DR Can Separate Mouse X and Y Sperm by Specific Inhibition of X Sperm Motility

The findings of Umehara [17] and Renfa [19] have shown that, after sperm treatment with R848, highly motile Y-bearing sperm are primarily located in the upper layer, while the lower layer predominantly contains slow-moving X-bearing sperm. To investigate whether the treatment of sperm with dsRNA-40 or dsRNA-DR can alter the ratio of X/Y sperm in the upper and lower layers, this study employed quantitative PCR based on the X-chromosome-specific gene *Akap4* and the Y-chromosome-specific gene *ddx3y* to assess the percentage of Y and X sperm in the upper and lower layers.

The results indicate that, in the 0.3 μ M dsRNA-40 treatment group, the percentage of Y sperm in the upper layer was $71.59\% \pm 3.73\%$ (Figure 3a), while the percentage of X sperm in the lower layer was $79.48\% \pm 1.44\%$ (Figure 3b). This study employed the dsRNAs transfection of spermatozoa in 1 mL of HTF medium, distributed into 1.5 mL centrifuge tubes. The upper layer was defined as the top 350 μ L of sperm resuspension in the centrifuge tubes, while the lower layer was designated as the bottom 350 μ L of sperm resuspension, mouse sperm collection methods and a schematic are provided in Supplementary Material Figure S1. In the 0.3 μ M dsRNA-DR treatment group, the percentage of Y sperm in the upper layer was $68.17\% \pm 2.72\%$ (Figure 3a), and the percentage of X sperm in the lower layer was $60.43\% \pm 23.07\%$ (Figure 3b). These findings suggest that the two tested dsRNAs can separate mouse X and Y sperm into two different layers. CASA analysis results indicate that the two dsRNA treatments significantly decreased the motility of the X-sperm-containing lower-layer sperm, but not the Y-sperm-containing upper-layer sperm, as compared to the control group (Figure 3c). The combined results of quantitative PCR and CASA analysis suggest that dsRNA-40 and dsRNA-DR can separate mouse X and Y sperm by the specific inhibition of X sperm motility, thereby leading to the upper layer containing predominantly high-motility Y sperm, while the lower layer predominantly consisting of slow-moving X sperm.

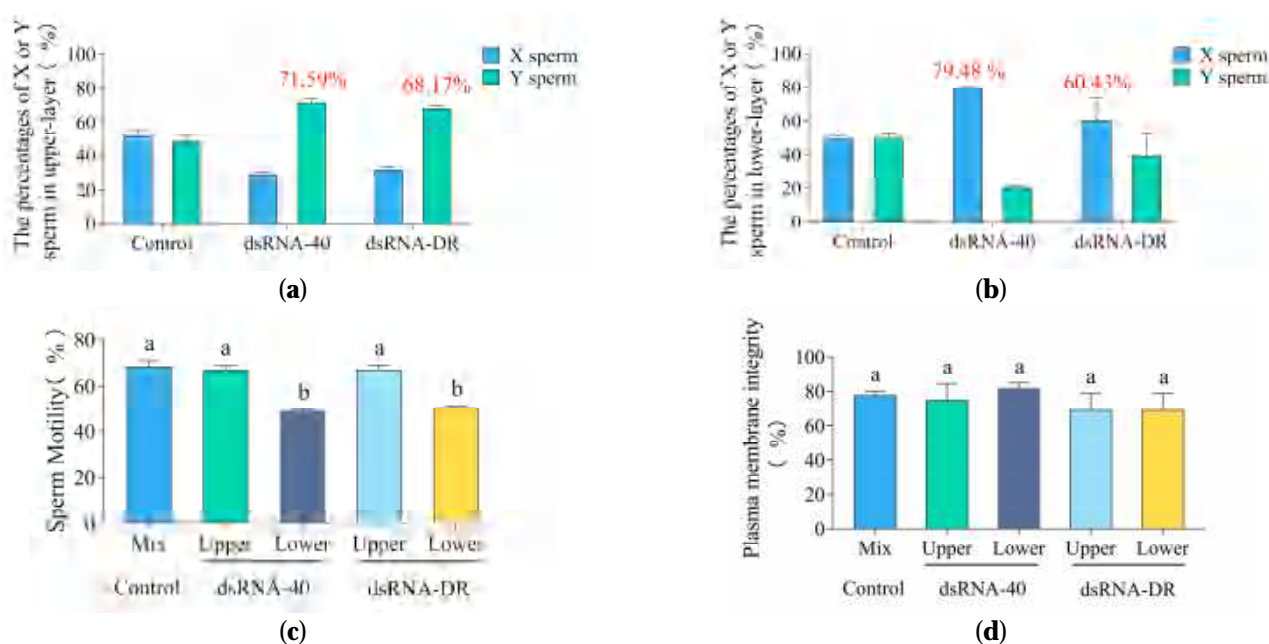


Figure 3. Cont.

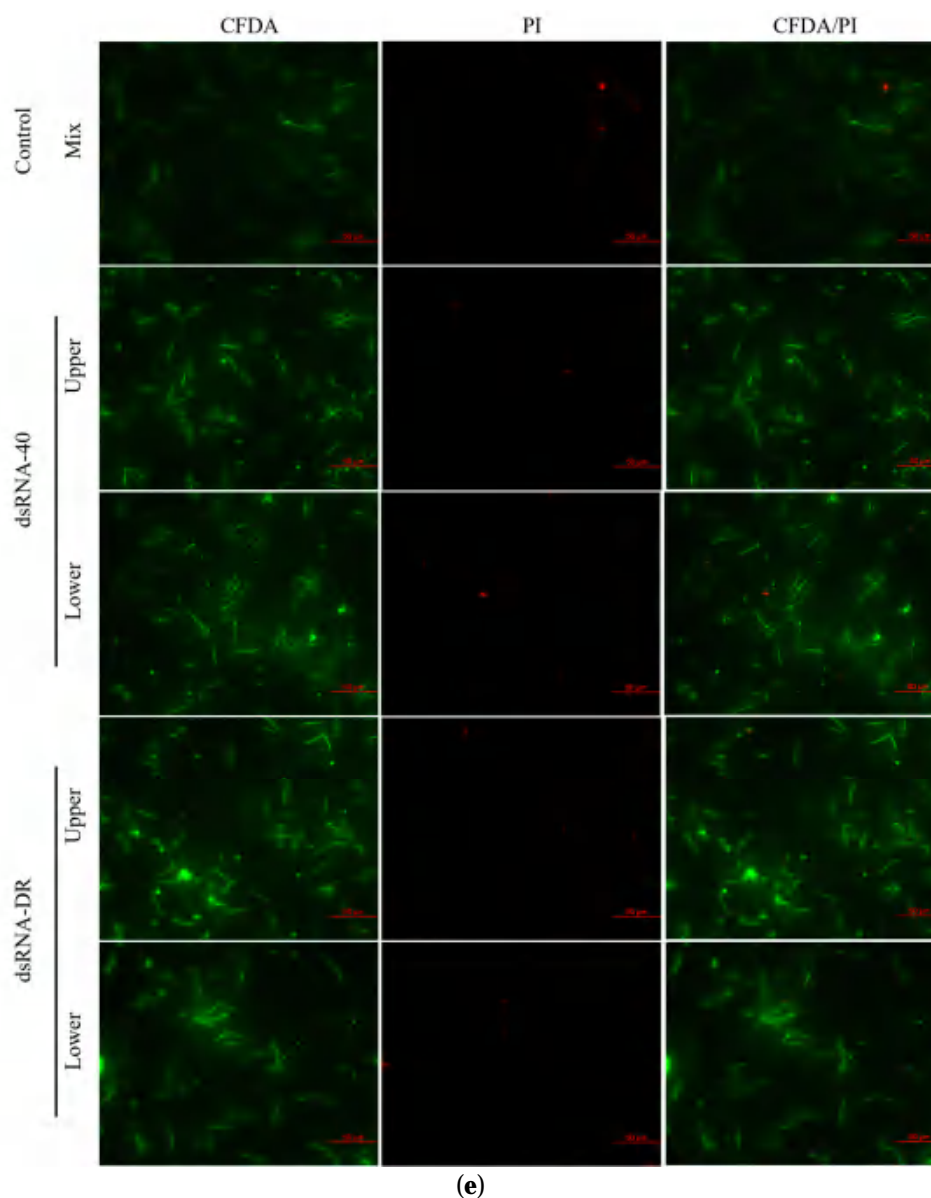


Figure 3. dsRNA-40 and dsRNA-DR can separate mouse X and Y sperm by the specific inhibition of X sperm motility. (a) The percentage of X and Y sperm in the upper layer after incubation with 0.3 μ M dsRNA-40 or dsRNA-DR for 30 min; (b) the percentage of X and Y sperm in the lower layer after incubation with 0.3 μ M dsRNA-40 or dsRNA-DR for 30 min; (c) comparison of motility in the upper and lower layers of sperm after incubation with 0.3 μ M dsRNA-40 or dsRNA-DR for 30 min; (d) comparison of plasma membrane integrity of sperm in the upper and lower layers after incubation with 0.3 μ M dsRNA-40 or dsRNA-DR for 30 min; (e) sperm plasma membrane integrity was analyzed by staining with 6-CFDA/PI. Different letters indicate statistically significant differences between groups, $p < 0.05$, while the same letters indicate no significant differences between groups, $p > 0.05$.

To further investigate whether the difference in sperm motility between the upper and lower layers is a result of sperm damage, we employed the 6-CFDA/PI assay kit to assess sperm plasma membrane integrity. The results in Figure 3d,e demonstrate that the plasma membrane integrity of both the upper- and lower-layer sperm is not significantly reduced compared to the control group. This finding suggests that dsRNA-40 or dsRNA-DR treatment has no significant impact on sperm plasma membrane integrity.

2.4. dsRNA-40 and dsRNA-DR Inhibit X Sperm Motility by Reducing ATP Level and Mitochondrial Activity

Sperm velocity is regulated by intracellular ATP levels, with mitochondria serving as a crucial energy source for sperm motility. To investigate whether the differences in sperm motility between the upper and lower layers result from variations in mitochondrial activity and ATP levels, we assessed the ATP levels and mitochondrial activity of sperm after treatment with dsRNA-40 or dsRNA-DR.

Sperm ATP levels were determined using the standard curve formula. The high linearity coefficient (R^2) of the ATP standard curve, reaching 0.9993, indicates its remarkable reliability (Figure 4a). As depicted in Figure 4b, after incubation with two tested dsRNAs for more than 10 min, sperm ATP levels began to decline compared to the control group, and this decline trend persisted over time, indicating that both dsRNA-40 and dsRNA-DR could inhibit sperm ATP production.

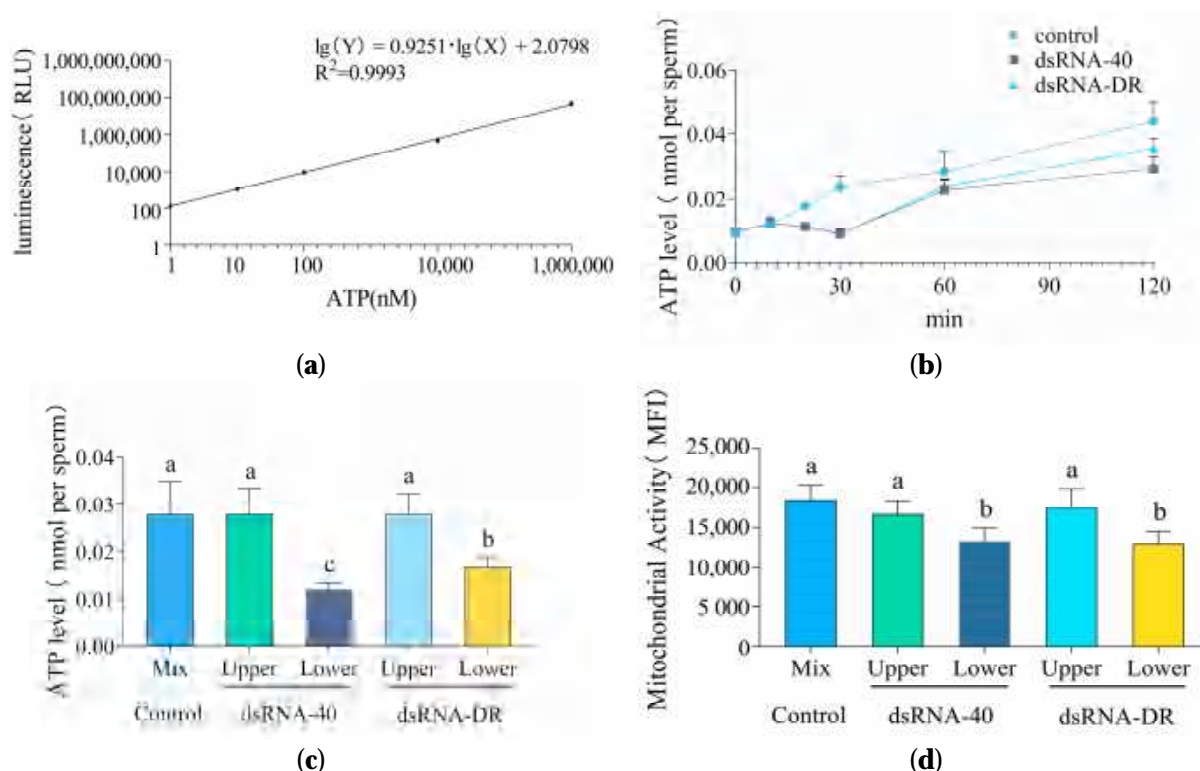


Figure 4. dsRNA-40 and dsRNA-DR inhibit X sperm motility by reducing ATP level and mitochondrial activity. (a) ATP standard curve with ATP concentrations of 1 mM, 10 μ M, 100 nM, 10 nM, and 1 nM; (b) changes in sperm ATP levels following incubation with or without 0.3 μ M dsRNA-40 or dsRNA-DR; (c) comparison of ATP levels between the upper and lower layer sperm after treatment with 0.3 μ M dsRNA-40 or dsRNA-DR for 30 min; (d) comparison of mitochondrial activity between the upper and lower layer sperm after treatment with 0.3 μ M dsRNA-40 or dsRNA-DR for 30 min. Different letters indicate statistically significant differences between groups, $p < 0.05$, while the same letters indicate no significant differences between groups, $p > 0.05$.

Following treatment with 0.3 μ M dsRNA-40 or dsRNA-DR for 30 min, a decrease in ATP levels was observed, primarily in the lower layer of sperm with reduced motility (Figure 4c). Moreover, compared to the control group, mitochondrial activity in the lower-layer sperm was significantly reduced (Figure 4d). These data confirm that dsRNA-40 and dsRNA-DR impact X sperm motility by reducing ATP levels and mitochondrial activity.

2.5. Production of Embryos with a Male-Biased Sex Ratio by Using Y Sperm Preselected with dsRNA-40 and dsRNA-DR Treatment for IVF

In this study, we developed a reliable method for determining the sex of individual blastocysts using a dual-PCR-based method (Figure 5a). To further substantiate that the upper layer of highly motile sperm predominantly consists of Y sperm, in vitro fertilization was conducted using the upper layer of mouse sperm separated by dsRNA treatment. The results displayed in Figure 5b show that when in vitro fertilization was carried out using the upper-layer mouse sperm preselected by dsRNA-40 or dsRNA-DR incubation, a total of 96, 184, and 109 embryos were obtained. Among these, male (XY) embryos accounted for 48 (50.17% \pm 8.42%), 138 (74.93% \pm 1.68%), and 73 (65.90% \pm 10.70%), while female (XX) embryos accounted for 48 (49.83% \pm 8.42%), 46 (25.06% \pm 1.68%), and 36 (34.10% \pm 10.70%), in the control, dsRNA-40 and dsRNA-DR groups, respectively. The production of predominantly male IVF embryos from the upper-layer sperm indicates the successful separation of mouse X and Y sperm by dsRNA treatment.

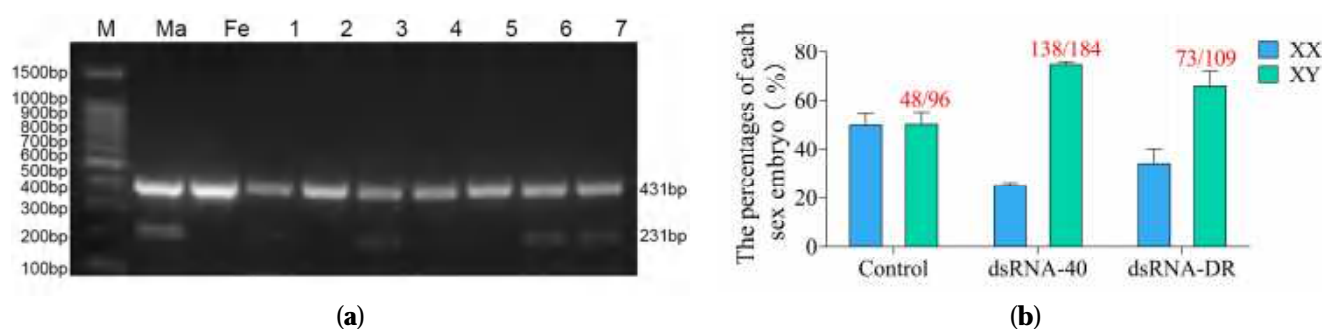


Figure 5. Production of embryos with a male-biased sex ratio by using Y sperm preselected with dsRNA-40 and dsRNA-DR treatment for IVF. (a) Dual PCR-based sex determination of in vitro-fertilized embryos (blastocysts) generated by the upper layer of sperm separated with 0.3 μ M dsRNA-40 or dsRNA-DR treatment for 30 min. To determine the gender of individual mouse blastocysts and to prevent false negatives due to sample loss, a dual PCR approach was employed. The upper band (431 bp) indicates the presence of an autosome-linked gene, while the lower band (231 bp) signifies the presence of a Y-chromosome-linked gene. The upper band is used to verify the success of PCR, and the lower band is used to determine the sexuality of tested IVF embryos. The male control group successfully amplified a band corresponding to the Y-chromosome-specific gene *ddx3y*. On the other hand, the female control group failed to produce an amplification band of such an amount. These findings demonstrate the accuracy of the dual PCR method for embryo sex identification. M, male control. F, female control. #1, #3, #6 and #7 were identified as male embryos, while #2, #4 and #5 were identified as female embryos; (b) the gender ratio of IVF embryos produced by the upper layer of sperm separated with 0.3 μ M dsRNA-40 or dsRNA-DR treatment for 30 min. Three experiments were conducted in each group. The total number of identified male embryos and the total number of analyzed IVF embryos are shown at the top of each group's column.

3. Discussion

TLR7/8 are toll-like receptors (TLRs) that perceive nucleic acids. They are usually located in the lysosomes within the cells. Segregated TLR7/8 TLR effectively prevents their contact with host DNA or RNA, thus avoiding the activation of innate immune responses. This also determines that extracellular RNA or DNA need to be transported into cells for the binding of nucleic-acid-sensing TLRs [30]. The TLR7/8 agonist R848 is a small-molecule compound, which can easily penetrate the cell membrane to activate TLR7/8 [34,35]. However, without cholesterol modifications, dsRNA-40 and dsRNA-DR could not pass through the sperm cell membrane. To facilitate the successful intracellular activation of TLR7/8 by dsRNA-40 and dsRNA-DR, we introduced cholesterol modifications at the 3' end of the dsRNAs. While significantly increasing the transfection efficiency of dsRNAs, cholesterol modifications seem not to alter the binding function of dsRNA-40 and dsRNA-

DR to TLR7/8, as the motility of TLR7/8-expressing X sperm was specifically inhibited by dsRNA-40 and dsRNA-DR. A similar result was also reported in another study [36].

Previous studies [17,19] indicated that the binding of R848 agonist to receptors TLR7/8 decreases ATP production and mitochondrial function by regulating the downstream GSK3 α/β -hexokinase pathway specifically in X sperm, without impacting Y sperm. We found in this study that the incubation of mouse sperm with TLR7/8 ligands dsRNA-40 and dsRNA-DR also reduces ATP level and mitochondrial activity in X sperm. This suggests that dsRNA-40 and dsRNA-DR also act through the downstream GSK3 α/β -hexokinase pathway of TLR7/8 to suppress the motility of X sperm.

In this study, the treatment of mouse sperm with 0.3 μ M dsRNA-40 or dsRNA-DR for 30 min resulted in the upper layer of sperm carrying 68.17–71.59% of Y sperm. However, the incubation of mouse sperm with 0.3 μ M R848 for 30 min resulted in 91.0% of sperm in the upper layer being Y sperm [17]. Furthermore, the use of dsRNA-40- or dsRNA-DR-preselected upper layer of sperm for IVF-generated mouse embryos resulted in a male ratio of 65.9–74.93%, while the use of R848-separated upper layer of sperm for IVF-produced mouse embryos resulted in a male ratio of 89.3% [17]. The effectiveness of the two investigated dsRNAs on separating mouse X and Y sperm seems lower than that of R848. This difference could be related to the difference in TLR7/8 binding activity between dsRNAs and R848 because dsRNAs are weaker TLR7/8 agonists than R848 [37–39]. The difference in X and Y sperm separation efficiency between dsRNAs and R848 also could be due to their difference in half-life, since dsRNAs are easily degraded [40,41], whereas R848 is a stable compound [22]. However, the efficiency of dsRNA-40 or dsRNA-DR in separating mouse X and Y sperm could be further improved by optimizing the incubation concentration and time with mouse sperm.

When choosing TLR7/8 agonists to separate X and Y sperm, the cost of the agonists is also a critical factor that should be considered, in addition to the effectiveness. Some company websites (InvivoGen Company, San Diego, CA, USA, <https://www.invivogen.com/r848>, accessed on 1 November 2023; Miltenyi Biotec, Bergisch Gladbach, Germany, <https://www.miltenyibiotec.com/CN-en/products/r848-resiquimod.html>, accessed on 1 November 2023) indicate that the price of 500 μ g (1.43 μ Moles) of R848 is 2087 CNY, which is approximately equivalent to 560 OD of dsRNA. The synthesis of this amount of dsRNA by some company (GenePharma Company, Suzhou, China, website: <http://www.genepharma.cn/>, accessed on 1 November 2023) costs about 9000 CNY. Although the dsRNA price looks higher than R848, it is a specialized custom product, and its cost is expected to significantly decrease with large-scale synthesis.

Although we demonstrated that tge dsRNA-40 and dsRNA-DR treatment of mouse sperm is feasible in the gender control of embryos, whether they can be employed to produce offspring with a skewed sex ratio is unknown. In addition, due to the lack of specialized equipment and knowledge of mouse embryo transfer technique in our laboratory, we did not use this technique to further validate the role of dsRNAs, which is a major drawback of this study. Since artificial insemination techniques are not particularly mature in mice, we did not use dsRNA-treated mouse sperm for artificial insemination to produce progeny in the present study. Nevertheless, artificial insemination techniques are easily performed and usually have a high success rate in livestock such as cattle and pig. Future studies can be conducted to test whether dsRNA-40 and dsRNA-DR are feasible in the sex ratio control of livestock.

4. Materials and Methods

4.1. Animals

Eight-week-old male ICR mice (Supplier: SPF (Beijing) Biotechnology Co., Ltd., Beijing, China, Catalog No. YD012) and eight-week-old female ICR mice (Supplier: SPF (Beijing) Biotechnology Co., Ltd., Catalog No. YD012) were utilized in this study. The use of experimental animals was approved by the Ethics Committee of the Experimental Animal Center of South China Agricultural University (License No.: SYXK-2019-0136).

4.2. Biochemical Materials

RNA was custom-designed and synthesized by Suzhou GenePharma Co., Ltd. (Suzhou, China). The carboxyfluorescein (FAM) moiety was added at the 5' end of dsRNA, and cholesterol modification was added at the 3' end of dsRNA. dsRNA-40 sense (5'-3'): GC-CCGUCUGUUGUGUGACUCTT; antisense (5'-3'): GAGUCACACAACAGACGGGCTT. DsRNA-DR sense (5'-3'): UGACCUUCA AUGUCCUUCAATT; antisense (5'-3'): UUGAAG-GACAUUGAAGGUCATT. The qPCR primers were designed using Primer Premier 6.

Human chorionic gonadotropin hCG (Catalog No.: M2520) and pregnant mare serum gonadotropin PMSG (Catalog No.: M2620), HTF fertilization medium (Catalog No.: M1130), M2 culture medium (Catalog No.: M1250), KSOM embryo culture medium (Catalog No.: M1430), and hyaluronidase (Catalog No.: M2215) were all obtained from Nanjing Aibei Biotechnology Co., Ltd. (Nanjing, China).

PCR primers and the direct PCR amplification kit (B639289) were procured from Sangon Biotech Co., Ltd. (Shanghai, China).

PowerUp SYBR Green Master Mix (A25742) was obtained from Thermo Fisher Scientific (Shanghai, China).

The ATP cell viability assay kit (Catalog No.: 40210ES10) was sourced from Yisheng Biotechnology Co., Ltd. (Shanghai, China).

The mitochondrial membrane potential assay kit (TMRE) (Catalog No.: C2001S) was obtained from Shanghai Biyuntian Biotechnology Co., Ltd. (Shanghai, China).

All other materials not explicitly mentioned were purchased from Sigma-Aldrich China (Shanghai, China).

4.3. Sperm Transfection with Cholesterol-Modified dsRNAs

The mouse spermatozoa were co-incubated with dsRNA-40, dsRNA-DR with or without cholesterol modification, at a concentration of 0.3 μ M for 10 min at 37 °C, in a 5% CO₂ incubator. The fluorescent probe FAM (carboxyfluorescein) was used with an absorption wavelength of 492 nm and an emission wavelength of 518 nm. A fluorescence microscope was used to detect green fluorescence production in sperm.

4.4. Assessment of Sperm Motility Using the CASA System

Sperm were incubated at 37 °C in an HTF medium containing 0.3 μ M of dsRNA-40 or dsRNA-DR for 30 min. Subsequently, the motility parameters of sperm in the upper and lower layers were assessed using a computer-assisted sperm analysis (CASA) system (HVIEW-SSAV600.8, Fuzhou Hongshiye Software Technology Co., Ltd., Fuzhou, China). Each experiment was repeated four times.

4.5. Measurement of Sperm Plasma Membrane and Mitochondrial Activity

The integrity of the plasma membrane of mouse sperm treated with dsRNA-40 or dsRNA-DR was assessed using the 6-carboxyfluorescein diacetate/propidium iodide (6-CFDA/PI) assay kit in HTF medium. Initially, 10 μ L of 6-CFDA was gently mixed with the treated mouse sperm and incubated at 37 °C for 30 min. Subsequently, 5 μ L of PI reagent was added, and incubation was continued for 10 min at 37 °C. Double-stained sperm (20 μ L) were transferred to a microscope slide, covered with a coverslip, and immediately observed and counted at 200 \times magnification under a fluorescence microscope, ensuring that no fewer than 200 sperm were observed in each field of view. The 6-CFDA freely diffuses inside cells and is hydrolyzed by non-specific lipase within the cell to produce carboxyl fluorescein (CF). The fluorescence product accumulates only in cells with intact cell membranes and active esterase activity, while dead cells remain unstained.

To assess changes in mitochondrial membrane potential in mouse sperm following treatment with dsRNA-40 or dsRNA-DR, 10 μ L of TMRE staining working solution was gently mixed with the treated mouse sperm and incubated for 30 min at 37 °C. Subsequently, centrifugation at 1000 \times g for 3 min was performed, followed by two washes. The measurements were carried out using a multi-mode microplate reader (BioTek Synergy H1,

Agilent Technologies, Santa Clara, CA, USA). TMRE is an orange-red cationic fluorescent probe that permeates the cell membrane, accumulates in intact mitochondria, and decreases in depolarized or inactive mitochondria, resulting in reduced TMRE accumulation. The excitation/emission wavelengths were 550 nm/575 nm.

4.6. Determination of Sperm ATP Content

Sperm ATP content measurement was conducted using the ATP assay kit (Catalog No.: 40210ES10, Yeasen Biotechnology, Shanghai, China). Sperm were incubated in HTF medium for various durations (0, 10, 20, 30, 60, and 120 min), both with and without treatment with dsRNA-40 or dsRNA-DR. Equal aliquots of samples from the upper and lower layers (100 μ L, diluted to 1×10^7 sperm/mL) were added to a 96-well white polystyrene microplate, and 100 μ L of the assay reagent was added to each well. The mixture was shaken at room temperature for 2 min to facilitate cell lysis and then left at room temperature for 10 min until stabilization. Chemical luminescence measurements were subsequently performed using a multi-mode microplate reader (BioTek Synergy H1, Agilent Technologies) with a detection wavelength of 560 nm.

An ATP standard curve was generated by adding various concentrations of ATP standard solutions (1 mM, 10 μ M, 100 nM, 10 nM, 1 nM) to each well of the 96-well plate, with each well containing 100 μ L of the standard solution. The ATP content was calculated based on the ATP standard curve to determine the relative vitality of the cells.

4.7. Measurement of the Percentage of X and Y Sperm in the Upper and Lower Layer of dsRNA-Treated Mouse Sperm

Sperm were incubated in an HTF medium containing 0.3 μ M of dsRNA-40 or dsRNA-DR at 37 °C for 30 min. Subsequently, DNA was extracted from 350 μ L of culture fluid collected from the upper layer, and the X/Y ratio of spermatozoa was determined by quantitative PCR.

The qPCR primers were designed to detect X and Y mouse sperm using the following sequences: Akap4-Forward: GGTAAGTAGCCACGCAAACTC, Akap4-Reverse: CATTGAGGAGCCAGTTGAGGACAC (X sperm, PCR product size 134 bp); Ddx3y-Forward: TCAGAAAAGCGTAACAACTGGG, Ddx3y-Reverse: AGGACAACTTGTGAGGGGAGC (Y sperm, PCR product size 161 bp); Gapdh-Forward: ATCTGAAAGACAAGAAACAGGGG, Gapdh-Reverse: TTGTGGTACGTGCATAGCTGA (Internal reference gene, PCR product size 148 bp). PCR was conducted using a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems, Waltham, MA, USA) at 95 °C for 2 min, followed by 40 cycles consisting of denaturation at 95 °C for 15 s, annealing at 58.5 °C for 30 s, and extension at 72 °C for 20 s.

The Power SYBR Green master mix (Thermo Fisher) was used according to the manufacturer's instructions. Each primer was confirmed to produce a single, specific PCR product through a melting curve program.

4.8. Identification of the Gender of IVF Embryos Generated by the Upper Layer of dsRNA-Treated Mouse Sperm

Sperm were incubated in an HTF medium containing 0.3 μ M of dsRNA-40 or dsRNA-DR at 37 °C for 30 min. Then, 350 μ L of sperm from the upper layer was collected and transferred to a new centrifuge tube. After centrifuging at 700 g for 4 min, the supernatant was discarded. The sperm were washed twice with 400 μ L of overnight-equilibrated HTF medium and then resuspended in 50 μ L of HTF medium and placed in the incubator.

Each 8-week-old female mouse was injected with 10 IU of PMSG at 18:00, followed by an injection of 10 IU of hCG 48 h later. After 14 h from the hCG injection, cumulus-oocyte complexes (COC) were collected from the female mouse's oviduct. For each in vitro fertilization (IVF) experiment, 25 COCs were placed in 50 μ L of HTF culture medium.

Five microliters of resuspended sperm were added to the fertilization drop, and the sperm density was adjusted to $1\text{--}5 \times 10^6$ /mL. Six hours after fertilization, the number of pronuclei in the oocyte was assessed to evaluate fertilization, followed by further culture in KSOM for embryo development analysis.

Some embryos were collected 4.5 days later, and PCR was employed for sexuality identification. The PCR primers were created to detect both Y chromosome and autosomes by following the given sequences: Rps18-Forward: TCTGTGGAGCGATGGAGTGTGA, Rps18-Reverse: GGCAGTGATGGCGAAGGCTATT (autosomes, PCR product size of 431 bp); Ddx3y-Forward: TGGACCAGCAAGTAAGTTGAACCT, Ddx3y-Reverse: GCCATTATGCGAAGCCTGTAGAGA (Y chromosome, PCR product size 231 bp). Each blastocyst was transferred to a 0.2 mL microcentrifuge tube containing 2 μ L of KSOM embryo culture medium and subjected to PCR amplification using the direct PCR amplification kit (Sangon Biotech) as per the manufacturer's instructions. ProFlex 3 \times 32 well PCR System 9700 (Applied Biosystems, USA) was used with 1 min incubation at 98 $^{\circ}$ C, followed by 40 cycles of denaturation at 98 $^{\circ}$ C for 5 s, annealing at 64 $^{\circ}$ C for 5 s, and extension at 72 $^{\circ}$ C for 20 s. The PCR products were then visualized on 2% (*w/v*) agarose gels.

4.9. Statistical Analysis

All data are presented as the mean \pm standard error from at least three independent experiments. Before statistical analysis, all data were tested for homogeneity of variance. Then, Student's *t*-test or one-way ANOVA was used to evaluate the statistical significance between the two groups of data. Statistical analyses were conducted using GraphPad Prism 9.0 (GraphPad Software, San Diego, CA, USA). Data were subjected to one-way analysis of variance (ANOVA) and two-way ANOVA. Values with *p* < 0.05 were considered statistically significant.

5. Conclusions

TLR7/8 ligands dsRNA-40 and dsRNA-DR are effective in the separation of mouse X and Y sperm by the specific inhibition of X sperm motility by reducing its ATP level and mitochondrial activity. The use of a dsRNA-40- or dsRNA-DR-preselected upper layer of sperm for IVF produces embryos with a sex ratio skewed toward males. The dsRNA-40 and dsRNA-DR treatment of sperm is a potential method for mammalian animal gender control.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/molecules29010262/s1>, Figure S1: Technical Diagram of Mouse X/Y Sperm Isolation Method using dsRNA. Mouse spermatozoa were resuspended in 1 mL HTF containing 0.3 μ M dsRNA and incubated at 37 $^{\circ}$ C for 30 minutes. Spermatozoa from the upper layer (Y spermatozoa) and lower layer (X spermatozoa) were then analyzed.

Author Contributions: Conceptualization, Z.L., J.P. and Y.H.; methodology, Y.H.; validation, J.P. and Y.H.; formal analysis, Z.L. and Y.H.; resources, L.H., Z.W., E.Z. and Z.L.; writing—original draft preparation, Y.H.; writing—review and editing, Z.L., J.P. and Y.H.; writing—review and editing, Z.L. and Y.H.; visualization, J.P. and Y.H.; supervision, L.H., Z.W., E.Z. and Z.L.; funding acquisition, Z.L. and L.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Guangdong Key Research and Development Project, grant number 2021B0707010007; 2018B030313011 and was funded by the Double first-class discipline promotion project, grant number 2023B10564001.

Institutional Review Board Statement: The animal study protocol was approved by the Ethics Committee of the Experimental Animal Center of South China Agricultural University (License No: SYXK-2019-0136).

Informed Consent Statement: Informed consent was not applicable as there were no human subjects involved in this article.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author (Z.L.), upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Oishi, K.; Hirooka, H. Effects of Sex Control and Twinning on Economic Optimization of Culling Cows in Japanese Black Cow-Calf Production Systems. *Theriogenology* **2012**, *77*, 320–330. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, X.; Li, J.; Chen, S.; Yang, N.; Zheng, J. Overview of Avian Sex Reversal. *Int. J. Mol. Sci.* **2023**, *24*, 8284. [\[CrossRef\]](#) [\[PubMed\]](#)
- Khalajzadeh, S.; Nejati-Javaremi, A.; Yeganeh, H.M. Effect of Widespread and Limited Use of Sexed Semen on Genetic Progress and Reproductive Performance of Dairy Cows. *Animal* **2012**, *6*, 1398–1406. [\[CrossRef\]](#) [\[PubMed\]](#)
- Khrantsova, E.A.; Wilson, M.A.; Martin, J.; Winham, S.J.; He, K.Y.; Davis, L.K.; Stranger, B.E. Quality Control and Analytic Best Practices for Testing Genetic Models of Sex Differences in Large Populations. *Cell* **2023**, *186*, 2044–2061. [\[CrossRef\]](#) [\[PubMed\]](#)
- Khrantsova, E.A.; Davis, L.K.; Stranger, B.E. The Role of Sex in the Genomics of Human Complex Traits. *Nat. Rev. Genet.* **2019**, *20*, 173–190. [\[CrossRef\]](#) [\[PubMed\]](#)
- Morrell, J.M. Applications of Flow Cytometry to Artificial Insemination: A Review. *Vet. Rec.* **1991**, *129*, 375–378. [\[CrossRef\]](#) [\[PubMed\]](#)
- Buchanan, B.R.; Seidel, G.J.; McCue, P.M.; Schenk, J.L.; Herickhoff, L.A.; Squires, E.L. Insemination of Mares with Low Numbers of Either Unsexed or Sexed Spermatozoa. *Theriogenology* **2000**, *53*, 1333–1344. [\[CrossRef\]](#) [\[PubMed\]](#)
- Catt, S.L.; Catt, J.W.; Gomez, M.C.; Maxwell, W.M.; Evans, G. Birth of a Male Lamb Derived from an in Vitro Matured Oocyte Fertilised by Intracytoplasmic Injection of a Single Presumptive Male Sperm. *Vet. Rec.* **1996**, *139*, 494–495. [\[CrossRef\]](#)
- Cran, D.G.; Johnson, L.A.; Miller, N.G.; Cochrane, D.; Polge, C. Production of Bovine Calves Following Separation of X- and Y-Chromosome Bearing Sperm and in Vitro Fertilisation. *Vet. Rec.* **1993**, *132*, 40–41. [\[CrossRef\]](#)
- Johnson, L.A.; Rath, D.; Vazquez, J.M.; Maxwell, W.M.C.; Dobrinsky, J.R. Preselection of Sex of Offspring in Swine for Production: Current Status of the Process and Its Application. *Theriogenology* **2005**, *63*, 615–624. [\[CrossRef\]](#)
- Rath, D.; Ruiz, S.; Sieg, B. Birth of Female Piglets Following Intrauterine Insemination of a Sow Using Flow Cytometrically Sexed Boar Semen. *Vet. Rec.* **2003**, *152*, 400–401. [\[CrossRef\]](#) [\[PubMed\]](#)
- Tubman, L.M.; Brink, Z.; Suh, T.K.; Seidel, G.E. Characteristics of Calves Produced with Sperm Sexed by Flow Cytometry/Cell Sorting. *J. Anim. Sci.* **2004**, *82*, 1029–1036. [\[CrossRef\]](#) [\[PubMed\]](#)
- Xi, J.; Wang, X.; Zhang, Y.; Jia, B.; Li, C.; Wang, X.; Ying, R. Sex Control by Zfy siRNA in the Dairy Cattle. *Anim. Reprod. Sci.* **2018**, *200*, 1–6. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yosef, I.; Edry-Botzer, L.; Globus, R.; Shlomovitz, I.; Munitz, A.; Gerlic, M.; Qimron, U. A Genetic System for Biasing the Sex Ratio in Mice. *EMBO Rep.* **2019**, *20*, e48269. [\[CrossRef\]](#) [\[PubMed\]](#)
- Douglas, C.; Maciulyte, V.; Zohren, J.; Snell, D.M.; Mahadevaiah, S.K.; Ojarikre, O.A.; Ellis, P.; Turner, J. CRISPR-Cas9 Effectors Facilitate Generation of Single-Sex Litters and Sex-Specific Phenotypes. *Nat. Commun.* **2021**, *12*, 6926. [\[CrossRef\]](#) [\[PubMed\]](#)
- Bai, M.; Liang, D.; Cheng, Y.; Liu, G.; Wang, Q.; Li, J.; Wu, Y. Gonadal Mosaicism Mediated Female-Biased Gender Control in Mice. *Protein Cell* **2022**, *13*, 863–868. [\[CrossRef\]](#) [\[PubMed\]](#)
- Umehara, T.; Tsujita, N.; Shimada, M. Activation of Toll-like Receptor 7/8 Encoded by the X Chromosome Alters Sperm Motility and Provides a Novel Simple Technology for Sexing Sperm. *PLoS Biol.* **2019**, *17*, e3000398. [\[CrossRef\]](#)
- Takashi, U.; Natsumi, T.; Zhendong, Z.; Moeka, I.; Masayuki, S. A Simple Sperm-Sexing Method That Activates TLR7/8 on X Sperm for the Efficient Production of Sexed Mouse or Cattle Embryos. *Nat. Protoc.* **2020**, *15*, 2645–2667.
- Ren, F.; Xi, H.; Ren, Y.; Li, Y.; Wen, F.; Xian, M.; Zhao, M.; Zhu, D.; Wang, L.; Lei, A.; et al. TLR7/8 Signalling Affects X-Sperm Motility via the GSK3 α/β -Hexokinase Pathway for the Efficient Production of Sexed Dairy Goat Embryos. *J. Anim. Sci. Biotechnol.* **2021**, *12*, 89. [\[CrossRef\]](#)
- Huang, M.; Cao, X.Y.; He, Q.F.; Yang, H.W.; Chen, Y.Z.; Zhao, J.L.; Ma, H.W.; Kang, J.; Liu, J.; Quang, F.S. Alkaline Semen Diluent Combined with R848 for Separation and Enrichment of Dairy Goat X-Sperm. *J. Dairy. Sci.* **2022**, *105*, 10020–10032. [\[CrossRef\]](#)
- Wen, F.; Liu, W.; Li, Y.; Zou, Q.; Xian, M.; Han, S.; Zhang, H.; Liu, S.; Feng, X.; Hu, J. TLR7/8 Agonist (R848) Inhibit Bovine X Sperm Motility via PI3K/GSK3 α/β and PI3K/NF κ B Pathways. *Int. J. Biol. Macromol.* **2023**, *232*, 123485. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, Z.; Kuo, J.C.; Zhang, C.; Huang, Y.; Zhou, Z.; Lee, R.J. A Squalene-Based Nanoemulsion for Therapeutic Delivery of Resiquimod. *Pharmaceutics* **2021**, *13*, 2060. [\[CrossRef\]](#) [\[PubMed\]](#)
- Greulich, W.; Wagner, M.; Gaidt, M.M.; Stafford, C.; Cheng, Y.; Linder, A.; Carell, T.; Hornung, V. TLR8 Is a Sensor of RNase T2 Degradation Products. *Cell* **2019**, *179*, 1264–1275.e13. [\[CrossRef\]](#) [\[PubMed\]](#)
- Guan, C.; Chernyak, N.; Dominguez, D.; Cole, L.; Zhang, B.; Mirkin, C.A. RNA-Based Immunostimulatory Liposomal Spherical Nucleic Acids as Potent TLR7/8 Modulators. *Small* **2018**, *14*, e1803284. [\[CrossRef\]](#)
- Heil, F.; Hemmi, H.; Hochrein, H.; Ampenberger, F.; Kirschning, C.; Akira, S.; Lipford, G.; Wagner, H.; Bauer, S. Species-Specific Recognition of Single-Stranded RNA via Toll-like Receptor 7 and 8. *Science* **2004**, *303*, 1526–1529. [\[CrossRef\]](#)
- Obermann, H.L.; Lederbogen, I.I.; Steele, J.; Dorna, J.; Sander, L.E.; Engelhardt, K.; Bakowsky, U.; Kaufmann, A.; Bauer, S. RNA-Cholesterol Nanoparticles Function as Potent Immune Activators via TLR7 and TLR8. *Front. Immunol.* **2021**, *12*, 658895. [\[CrossRef\]](#)
- Ohto, U.; Tanji, H.; Shimizu, T. Structure and Function of Toll-like Receptor 8. *Microbes Infect.* **2014**, *16*, 273–282. [\[CrossRef\]](#)
- Ewald, S.E.; Engel, A.; Lee, J.; Wang, M.; Bogoy, M.; Barton, G.M. Nucleic Acid Recognition by Toll-like Receptors Is Coupled to Stepwise Processing by Cathepsins and Asparagine Endopeptidase. *J. Exp. Med.* **2011**, *208*, 643–651. [\[CrossRef\]](#)
- Kagan, J.C.; Barton, G.M. Emerging Principles Governing Signal Transduction by Pattern-Recognition Receptors. *Cold Spring Harb. Perspect. Biol.* **2014**, *7*, a016253. [\[CrossRef\]](#)

30. Moresco, E.M.; LaVine, D.; Beutler, B. Toll-like Receptors. *Curr. Biol.* **2011**, *21*, R488–R493. [[CrossRef](#)]
31. Méndez-Acevedo, K.M.; Valdes, V.J.; Asanov, A.; Vaca, L. A Novel Family of Mammalian Transmembrane Proteins Involved in Cholesterol Transport. *Sci. Rep.* **2017**, *7*, 7450. [[CrossRef](#)] [[PubMed](#)]
32. Salim, L.; McKim, C.; Desaulniers, J.P. Effective Carrier-Free Gene-Silencing Activity of Cholesterol-Modified siRNAs. *RSC Adv.* **2018**, *8*, 22963–22966. [[CrossRef](#)] [[PubMed](#)]
33. Sanitt, P.; Apiratikul, N.; Niyomtham, N.; Yingyongnarongkul, B.E.; Assavalapsakul, W.; Panyim, S.; Udomkit, A. Cholesterol-Based Cationic Liposome Increases dsRNA Protection of Yellow Head Virus Infection in *Penaeus Vannamei*. *J. Biotechnol.* **2016**, *228*, 95–102. [[CrossRef](#)] [[PubMed](#)]
34. Hemmi, H.; Kaisho, T.; Takeuchi, O.; Sato, S.; Sanjo, H.; Hoshino, K.; Horiuchi, T.; Tomizawa, H.; Takeda, K.; Akira, S. Small Anti-Viral Compounds Activate Immune Cells via the TLR7 MyD88-Dependent Signaling Pathway. *Nat. Immunol.* **2002**, *3*, 196–200. [[CrossRef](#)]
35. Jurk, M.; Heil, F.; Vollmer, J.; Schetter, C.; Krieg, A.M.; Wagner, H.; Lipford, G.; Bauer, S. Human TLR7 or TLR8 Independently Confer Responsiveness to the Antiviral Compound R-848. *Nat. Immunol.* **2002**, *3*, 499. [[CrossRef](#)]
36. Karikó, K.; Buckstein, M.; Ni, H.; Weissman, D. Suppression of RNA Recognition by Toll-like Receptors: The Impact of Nucleoside Modification and the Evolutionary Origin of RNA. *Immunity* **2005**, *23*, 165–175. [[CrossRef](#)]
37. Abdel-Mageed, A.M.; Isobe, N.; Yoshimura, Y. Effects of Virus-Associated Molecular Patterns on the Expression of Cathelicidins in the Hen Vagina. *J. Poult. Sci.* **2016**, *53*, 240–247. [[CrossRef](#)]
38. Cekaite, L.; Furset, G.; Hovig, E.; Sioud, M. Gene Expression Analysis in Blood Cells in Response to Unmodified and 2'-Modified siRNAs Reveals TLR-Dependent and Independent Effects. *J. Mol. Biol.* **2007**, *365*, 90–108. [[CrossRef](#)]
39. Hotz, C.; Treinies, M.; Mottas, I.; Rötzer, L.C.; Oberson, A.; Spagnuolo, L.; Perdicchio, M.; Spinetti, T.; Herbst, T.; Bourquin, C. Reprogramming of TLR7 Signaling Enhances Antitumor NK and Cytotoxic T Cell Responses. *Oncoimmunology* **2016**, *5*, e1232219. [[CrossRef](#)]
40. Elbashir, S.M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschl, T. Duplexes of 21-Nucleotide RNAs Mediate RNA Interference in Cultured Mammalian Cells. *Nature* **2001**, *411*, 494–498. [[CrossRef](#)]
41. Gallas, A.; Alexander, C.; Davies, M.C.; Puri, S.; Allen, S. Chemistry and Formulations for siRNA Therapeutics. *Chem. Soc. Rev.* **2013**, *42*, 7983–7997. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



OPEN

A Transcriptome Analysis Reveals that Hepatic Glycolysis and Lipid Synthesis Are Negatively Associated with Feed Efficiency in DLY Pigs

Cineng Xu^{1,2}, Xingwang Wang^{1,2}, Zhanwei Zhuang¹, Jie Wu¹, Shenping Zhou¹, Jianping Quan¹, Rongrong Ding¹, Yong Ye¹, Longlong Peng¹, Zhenfang Wu¹, Enqin Zheng¹✉ & Jie Yang¹✉

Feed efficiency (FE) is an important trait in the porcine industry. Therefore, understanding the molecular mechanisms of FE is vital for the improvement of this trait. In this study, 6 extreme high-FE and 6 low-FE pigs were selected from 225 Duroc × (Landrace × Yorkshire) (DLY) pigs for transcriptomic analysis. RNA-seq analysis was performed to determine differentially expressed genes (DEGs) in the liver tissues of the 12 individuals, and 507 DEGs were identified between high-FE pigs (HE- group) and low-FE pigs (LE- group). A gene ontology (GO) enrichment and pathway enrichment analysis were performed and revealed that glycolytic metabolism and lipid synthesis-related pathways were significantly enriched within DEGs; all of these DEGs were downregulated in the HE- group. Moreover, Weighted gene co-expression analysis (WGCNA) revealed that oxidative phosphorylation, thermogenesis, and energy metabolism-related pathways were negatively related to HE- group, which might result in lower energy consumption in higher efficiency pigs. These results implied that the higher FE in the HE- group may be attributed to a lower glycolytic, energy consumption and lipid synthesizing potential in the liver. Furthermore, our findings suggested that the inhibition of lipid synthesis and glucose metabolic activity in the liver may be strategies for improving the FE of DLY pigs.

Feed efficiency (FE) is an important economic trait in the porcine industry, as feed cost accounts for approximately 70% of the total production cost¹. Therefore, improving FE can greatly promote the economic benefits of pig production. The main indicators for measuring FE are residual feed intake (RFI) or feed conversion ratio (FCR). RFI is defined as the difference between the animal's actual feed intake and its predicted dry matter intake (DMI) based on production needs, to specifically capture the efficiency of feed use independent from production needs². A lower RFI value indicates a more efficient pig. FCR is the ratio of feed intake to the average daily gain (ADG) during a specified period. Compared with FCR, RFI can more accurately reflect the differences of FE in pigs with different body weights and growth rate². Therefore, RFI is preferred as the selection indicator. Because of the strong genetic correlation with RFI (R equals 0.76–0.99)³, FCR can be used as a reference indicator to verify the selection based on RFI, which can measure FE more intuitively.

To date, based on genome-wide association analysis (GWAS) of pigs, some SNPs and candidate genes that might affect FE have been identified. SNPs located on SSC7, SSC13, SSC14, and SSC17 and candidate genes, such as *MBD5*, *GTF3C5*, *HMGA2*, *PITX2*, and *MAP3K14*, were reported to be associated with FE in the previous studies^{4–7}. However, GWAS studies are limited to finding chromosomal regions or preselected genes that affect FE⁸ and finding candidate genes and pathways that affect FE is difficult. Instead, RNA-seq technology can quantitatively measure gene expression in individuals to screen differentially expressed genes (DEGs)^{9,10}. By analyzing the DEGs and related biological pathways, candidate genes and pathways that affect FE can be identified.

¹College of Animal Science and National Engineering Research Center for Breeding Swine Industry, South China Agricultural University, Guangdong, P.R. China. ²These authors contributed equally: Cineng Xu and Xingwang Wang.

✉e-mail: eqzheng@scau.edu.cn; jieyang2012@hotmail.com

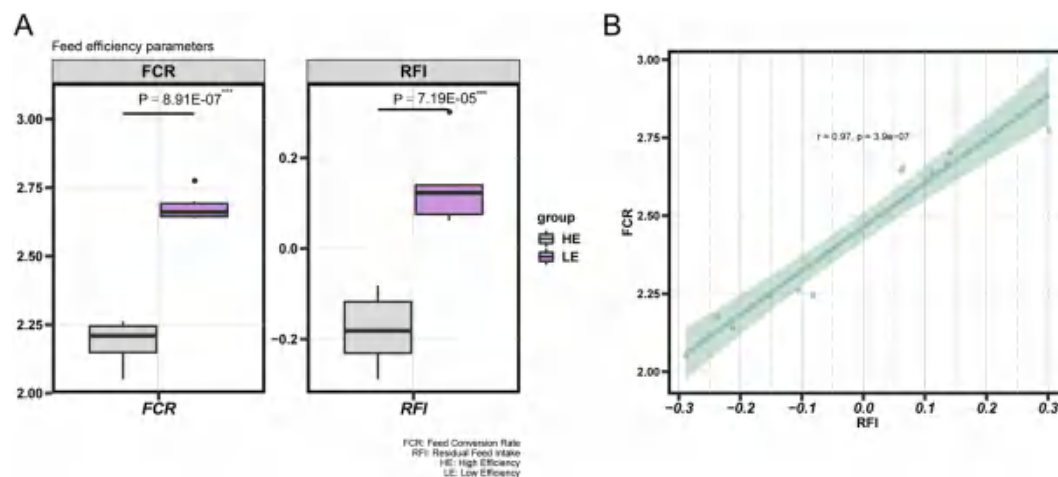


Figure 1. Feed efficiency (FE) phenotypic parameters in Duroc × (Landrace × Yorkshire) (DLY) pigs from high-FE pigs (HE- group) and low-FE pigs (LE- group). (A) Boxplot of the feed conversion ratio (FCR) and residual feed intake (RFI) in the two groups. (B) The correlation coefficient between FCR and RFI in the two groups.

RNA-seq has been used to research FE in animals, including pigs, cattle, and poultry, and muscle, liver, and adipose tissues have been used as research materials^{11–13}. Recently, a growing number of transcriptome analysis has focused on the liver tissue of pigs to identify candidate genes and pathways associated with FE^{14–16}. In mammals, the liver plays a prominent and central role in regulating the metabolism and distribution of nutrients. On the one hand, macronutrients, such as carbohydrates, lipids, and proteins, are metabolized in the liver^{17,18}. On the other hand, the liver can synthesize and store nutrients, as well as release nutrients into the blood¹⁹. Several studies have revealed that lipid metabolism, such as fatty acid synthesis, lipogenesis, and steroidogenesis, in the liver tissue was altered in FE-divergent pigs^{14,20,21}. In addition, glucose metabolism and energy metabolism in the liver have been reported to be essential for the regulation of FE in pigs, and lower glycolytic potential and energy loss were found in high-FE pigs¹⁵.

Many transcriptome studies focused on purebred pigs (including Yorkshire and Duroc) and crossbred pigs, such as Large White × (Landrace × Pietrain) and Maxgro × (Landrace × Large White) pigs, have made some progress in identifying candidate genes and pathways linked with FE^{14,20–22}. However, none of the transcriptome studies have been conducted in the liver of commercial Duroc × (Landrace × Yorkshire) (DLY) pigs, while DLY pigs are by far the largest population in the porcine industry worldwide²³. As a result, in this study, we used RNA-seq technology to profile the liver transcriptome of 6 extremely high-FE DLY pigs (HE- group) and 6 extremely low-FE DLY pigs (LE- group) to identify candidate genes and pathways that significantly correlated with the FE of DLY pigs. Furthermore, the identified genes and pathways that affect FE can provide theoretical support for pig selection, to improve the feed efficiency and economic benefits in commercial pig production in the future.

Results

Phenotypic parameters in Pigs from the HE- and LE- groups. Six extremely high-FE pigs had the phenotypic parameter of $RFI = -0.18 \pm 0.08$ and $FCR = 2.19 \pm 0.08$, while six extremely low-FE pigs had the phenotypic parameter of $RFI = 0.14 \pm 0.09$ and $FCR = 2.68 \pm 0.05$. The phenotypic details of the HE- and LE- groups are shown in Table S1. The FCR and RFI of the HE- group were significantly lower than those of the LE- group, which are displayed in the boxplot (Fig. 1A); thus, the HE- group was more efficient than the LE- group. A high linear positive correlation between FCR and RFI (Fig. 1B) was identified in our study, which was consistent with previous studies³.

Mapping statistics summary. In this study, the mapping statistics summaries for each sample are listed in Table S2. On average, the sequencing generated 40719764 and 40985844 effective reads in the HE- and LE- groups, respectively. Among the effective reads, on average, 93.36% and 92.81% in the HE- and LE- groups, respectively, were uniquely mapped to the reference genome, and 3.62% and 3.69% were multiple mapped to the reference genome.

A total of 507 DEGs between the HE- and LE- groups. In the current study, a total of 507 DEGs satisfied the criteria of $|\log_2(\text{Foldchange})| > 1$ and $q\text{-value} < 0.001$. Among the 507 DEGs, 53 DEGs were upregulated and 454 DEGs were downregulated in the HE- group; the 5 most significantly upregulated named genes (including *NRN1*, *CXCL13*, *DLK1*, *PLB1*, and *LYPD6B*) and 5 most significantly downregulated named genes (including *ADAMTS19*, *TRPV6*, *NME8*, *ANHX*, and *LRRC71*) are marked (Fig. 2). The details of all DEGs with their $\log_2(-\text{Foldchange})$ and $q\text{-value}$ are listed in Table S3.

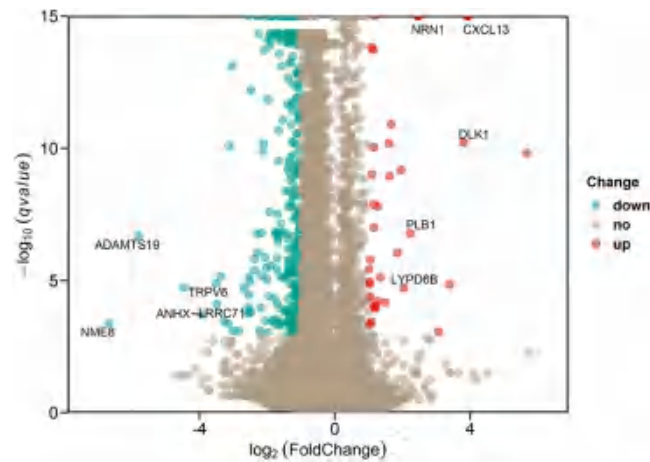


Figure 2. Differentially expressed genes (DEGs) between the HE- and LE- groups. A plot of DEGs with $|\log_2(\text{Foldchange})| > 1$ and $q\text{-value} < 0.001$. Red dots represent significantly upregulated genes, and green dots represent significantly downregulated genes. The genes marked in the figure are the 5 most significantly upregulated and 5 most downregulated named genes. The x-axis and y-axis represent the $-\log_{10}(q\text{value})$ and $\log_2(\text{Foldchange})$, respectively.

GO_ID	GO term	q-value	Gene Names
Biological process			
GO:0046835	carbohydrate phosphorylation	4.99E-03	HK3, PFKFB4, PFKFB1, GCK
Molecular function			
GO:0019200	carbohydrate kinase activity	1.01E-03	HK3, PFKFB4, PFKFB1, GCK
GO:0005088	Ras guanyl-nucleotide exchange factor activity	1.68E-03	ARHGEF37, ARHGEF39, ARHGEF16, ARHGEF25, DENND3
GO:0005089	Rho guanyl-nucleotide exchange factor activity	2.89E-03	ARHGEF39, ARHGEF16, ARHGEF37, ARHGEF25
GO:0005085	guanyl-nucleotide exchange factor activity	7.78E-03	ARHGEF37, ARHGEF39, ARHGEF16, ARHGEF25, DENND3

Table 1. Significantly enriched GO terms. Enriched liver DEGs between the HE- and LE- groups according to gene ontology (GO) terms for biological processes, molecular function, and cellular component. Significantly enriched terms ($q\text{-value} < 0.05$) are listed with the GO_ID, term, $q\text{-value}$ and gene name. The remaining terms are shown in Table S4.

GO enrichment analysis of DEGs. The biological process (BP), molecular function (MF), and cellular component (CC) GO terms of 507 DEGs were identified and the details of all identified GO terms are listed in Table S4. A total of 5 GO terms were significantly enriched ($q\text{-value} < 0.05$), including 1 BP_GO term and 4 MF_GO terms. All of the genes enriched in the 5 GO terms were downregulated in the HE- group (Table 1). None of the CC_GO terms was significantly enriched. The significantly enriched BP_GO term was carbohydrate phosphorylation, which is involved in glycolysis. The most significantly enriched MF_GO term was carbohydrate kinase activity. Carbohydrate kinase includes hexokinase, phosphofructokinase, and other kinases, which mainly catalyze glycolysis. Genes involved in the 2 terms were downregulated in the HE- group, indicating that glycolysis might decrease in the liver of the HE- group. The remaining 3 MF_GO terms were related to guanyl-nucleotide exchange factor activity, which exchanges GDP for GTP²⁴, and the DEGs enriched in these terms were downregulated in the HE- group.

Pathway enrichment analysis of DEGs. Our results showed that 25 significantly enriched pathways ($q\text{-value} < 0.05$) were enriched in the Reactome or KEGG database, 24 pathways were enriched in the Reactome database, and one pathway was enriched in the KEGG database (Table S5). The top 10 significantly enriched pathways and the genes contained in each pathway are listed in Table 2. Among the 10 pathways, 6 pathways were related to carbohydrate metabolism, including metabolism of carbohydrates, glucose transport, glycolysis, hexose transport, glucose metabolism, and starch and sucrose metabolism; the other 4 pathways were correlated with lipid synthesis, including lipid and lipoprotein metabolism, SREBP activation gene expression, SREBP regulation of cholesterol biosynthesis and phase 1 - functionalization of compounds. Most of the genes involved in the 10 pathways were downregulated in the HE- group. These results indicated that decreased lipid and glucose metabolism activity might occur in the liver of the HE- group.

Pathway	q-value	Gene Names
Reactome		
Metabolism of carbohydrates	5.52E-07	GCK, HK3, PFKFB1, PFKFB4, LOC100521789, RAE1, FGF21, LOC100739542, LOC100522672, CHST1
Glucose transport	4.82E-05	GCK, HK3, FGF21, RAE1, LOC100739542
Glycolysis	4.82E-05	GCK, HK3, PFKFB1, PFKFB4
Hexose transport	4.82E-05	GCK, HK3, FGF21, RAE1, LOC100739542
Glucose metabolism	0.001906	GCK, HK3, PFKFB1, PFKFB4
Metabolism of lipids and lipoproteins	0.0029006	ACACB, SCD, CYP21A2, CHKA, CYP7A1, HSD17B1, PLA2G4B, LSS, HMGCS2
Activation of gene expression by SREBP	0.0098664	ACACB, SCD, LSS
Regulation of cholesterol biosynthesis by SREBP	0.0099889	ACACB, SCD, LSS
Phase 1 - Functionalization of compounds	0.0124725	CYP21A2;AOC1;CYP7A1
KEGG pathway		
Starch and sucrose metabolism	0.039301	GCK, HK3, UGT1A1, LOC100516628, LOC100521789, GBE1, LOC100522672

Table 2. Top 10 significantly enriched pathways enriched in the Reactome and KEGG databases. Enriched DEGs according to the Reactome or Kyoto Encyclopedia of Genes and Genomes (KEGG) database are shown. Significantly enriched pathways (q -value < 0.05) are listed with the pathway, q -value and gene names. Twenty-four pathways were significantly enriched in the Reactome database, while only 1 pathway was significantly enriched in the KEGG database. The top 10 significantly enriched pathways are listed. The remaining pathways are shown in Table S5.

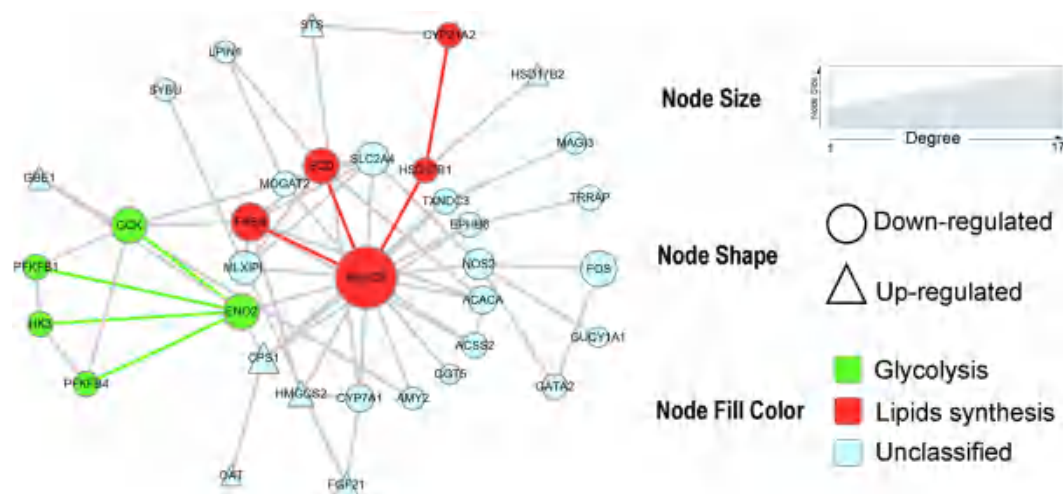


Figure 3. The key network of DEGs in the liver from HE- compared with LE- group. The network diagram centers on the ACACB gene, which has the largest degree of change, and the DEGs are directly or indirectly related to ACACB. Node shape represents the change in gene expression. The node fill color represents the functional classification of the gene.

Protein-protein interaction (PPI) analysis. According to Protein-protein interaction (PPI) analysis, a network of some of the named genes was visualized to explore the interaction of them with each other (Fig. 3). The network diagram was centered on the ACACB gene, which had the largest degree, and 32 DEGs were directly or indirectly related to it. Genes highlighted in green (GSK, HK3, ENO2, PFKFB1, and PFKFB4) are involved in glycolysis, and genes highlighted in red (ACACB, SCD, FASN, HSD17B1, and CYP21A2) are involved in lipid

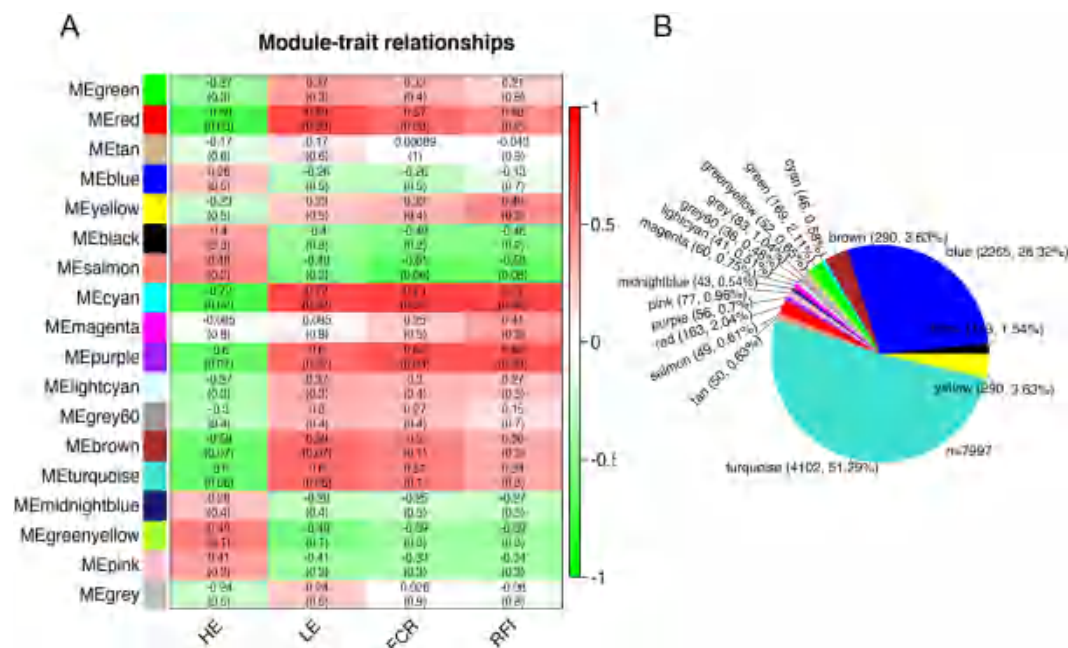


Figure 4. Weighted gene co-expression analysis (WGCNA). **(A)** Correlations between hepatic genes co-expression modules and feed efficiency traits of DLY pig. Modules represent the network of co-expressed genes and are named by different colors. Correlations are presented in the rectangles and the value in parentheses represent the p -value. **(B)** The number and percentage of genes in various modules.

synthesis. All genes highlighted in green or red color were downregulated in the HE- group, indicating that glycolysis and lipid synthesis might be reduced in the liver of the HE- group.

Weighted gene co-expression analysis (WGCNA) and enrichment pathways of modules correlated to feed efficiency traits. WGCNA was conducted to identify gene co-expression modules that are correlated with the trait of interest (HE, LE, RFI, and FCR). A total of 18 co-expressed gene modules were identified and named by different colors (Fig. 4A). The list of genes in these modules was presented in Table S6. Among them, two-thirds (12/18) are negatively correlated and one-third (6/18) are positively correlated with HE- group. This may imply that high feed efficiency may be associated with a lower level of certain biological processes. Two of these modules were significantly positively associated with RFI and FCR, namely the MEcyan module ($r = 0.72$, $p = 0.02$; $r = 0.74$, $p = 0.01$) and the MEpurple module ($r = 0.68$, $p = 0.03$; $r = 0.66$, $p = 0.04$). The MEcyan module ($r = -0.72$, $p = 0.02$) and MEred module ($r = -0.69$, $p = 0.03$) were negatively correlated with HE- group, while positively correlated with LE- group. The MEcyan module clustered 46 genes, and 56 genes were clustered in MEpurple module, while 163 in MEred (Fig. 4B).

The functional enrichment analysis for the three modules significantly correlated with FE-related traits was conducted in KEGG and Reactome database and were presented in Table S7. The MEcyan and MEpurple module identified 20 significantly enriched pathways in KEGG database, while 4 in Reactome database ($p.adjust < 0.05$) (Table 3). However, there was no significantly enriched pathway in MEred. Moreover, none of the significantly enriched terms were identified in GO analysis in the MEcyan, MEpurple and MEred modules (Table S8). Oxidative phosphorylation and thermogenesis were the most significantly enriched pathways identified in KEGG database, which contained 12 genes and related to energy consumption and negatively correlated with HE- group (Table 3). Three of four significantly enriched pathways enriched in Reactome database were correlated with energy metabolism, including “The citric acid (TCA) cycle and respiratory electron transport”, “Respiratory electron transport”, and “Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins” (Table 3).

Quantitative real-time PCR validation of six randomly selected DEGs. The reliability of the DEGs identified by RNA-seq was validated by qPCR. Six DEGs (*ACIN1*, *ACSS2*, *CCL21*, *HAMP*, *LSG1*, and *SAFB2*) were randomly selected for qPCR. Moreover, all of 12 individuals from the HE- and LE- groups were selected for qPCR. A significant correlation (P -value < 0.05) between the gene expression data calculated by RNA-seq and the gene expression data calculated by qPCR was found in 5 selected DEGs (*ACIN1*, *CCL21*, *HAMP*, *LSG1*, and *SAFB2*) (Fig. 5). Although the P -value of *ACSS2* was more than 0.05, it had a trend line similar to the other 5 selected DEGs. These results revealed a significant correlation between the two measures and confirmed the reliability of the gene expression data identified by RNA-seq.

Pathway	q-value	Gene Names
Reactome		
The citric acid (TCA) cycle and respiratory electron transport	0.020278616	NDUFV3, ADHFE1, NDUFB7, NDUFA11, NDUFA2
Complex I biogenesis	0.020278616	NDUFV3, NDUFB7, NDUFA11, NDUFA2
Respiratory electron transport	0.020278616	NDUFV3, NDUFB7, NDUFA11, NDUFA2
Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins.	0.042865302	NDUFV3, NDUFB7, NDUFA11, NDUFA2
KEGG pathway		
Oxidative phosphorylation	1.33E-09	NDUFV3, NDUFA13, UQCR10, NDUFB7, LOC100525869, LOC100516480, COX5B, COX17, NDUFA11, COX7A1, NDUFA2, ATP6V0C
Thermogenesis	3.18E-07	NDUFV3, LOC100516527, NDUFA13, UQCR10, NDUFB7, LOC100525869, LOC100516480, COX5B, COX17, NDUFA11, COX7A1, NDUFA2
Parkinson disease	3.18E-07	NDUFV3, NDUFA13, UQCR10, NDUFB7, LOC100525869, LOC100516480, COX5B, NDUFA11, COX7A1, NDUFA2
Non-alcoholic fatty liver disease (NAFLD)	3.72E-07	NDUFV3, NDUFA13, UQCR10, NDUFB7, LOC100525869, LOC100516480, COX5B, NDUFA11, COX7A1, NDUFA2
Huntington disease	3.72E-07	NDUFV3, NDUFA13, UQCR10, NDUFB7, LOC100525869, LOC100516480, COX5B, NDUFA11, COX7A1, NDUFA2, POLR2J
Alzheimer disease	7.96E-07	NDUFV3, NDUFA13, UQCR10, NDUFB7, LOC100525869, LOC100516480, COX5B, NDUFA11, COX7A1, NDUFA2
Cardiac muscle contraction	0.001847933	UQCR10, LOC100525869, LOC100516480, COX5B, COX7A1
Retrograde endocannabinoid signaling	0.002439495	NDUFV3, NDUFA13, NDUFB7, GNB2, NDUFA11, NDUFA2
Pentose and glucuronate interconversions	0.002655063	GUSB, LOC100516628, UGT2B31
Porphyrin and chlorophyll metabolism	0.007363959	GUSB, LOC100516628, UGT2B31
Sulfur metabolism	0.011099108	SELENBP1, TST
Drug metabolism - cytochrome P450	0.018779781	ADH4, LOC100516628, UGT2B31
Metabolism of xenobiotics by cytochrome P450	0.019524461	ADH4, LOC100516628, UGT2B31
Retinol metabolism	0.025126094	ADH4, LOC100516628, UGT2B31
Fatty acid metabolism	0.025126094	FADS1, FADS2, CBR4
Ascorbate and aldarate metabolism	0.025126094	LOC100516628, UGT2B31
Drug metabolism - other enzymes	0.025213695	GUSB, LOC100516628, UGT2B31
Chemical carcinogenesis	0.026132735	ADH4, LOC100516628, UGT2B31
Aminoacyl-tRNA biosynthesis	0.030796382	VARS1, HARS1, MARS1
Ribosome	0.033382135	RPL26L1, MRPL14, RPS21, RPL31

Table 3. A total of 20 significantly enriched pathways in KEGG database and 4 in Reactome database for MEcyan and MEpurple modules. Significantly enriched pathways (q -value < 0.05) are listed with the pathway, q -value and gene names.

Discussion

In this study, the DEGs, at the mRNA level, in the liver of the HE- and LE- groups were identified. Furthermore, the relevant pathways and candidate genes affecting the different FEs between the two groups were explored, illuminating the metabolic pathways and molecular mechanisms associated with the divergence in FE. We found that decreased glycolytic, energy consumption and lipid synthesizing potential in the liver may be associated with improved feed efficiency in DLY pigs.

The PPI analysis revealed some candidate genes associated with glycolysis (including *GCK*, *ENO2*, *HK3*, *PFKFB1*, and *PFKFB4*). Combining the results of GO and pathway enrichment analysis, *HK3*, *PFKFB1*, *GCK*, and *PFKFB4* were enriched in the most significantly enriched GO terms and pathways, all of which were downregulated in the HE- group. The *GCK* gene is involved in glycolysis and encodes hexokinase 4 that catalyzes glucose transfer to glucose-6-phosphate^{25,26}. The overexpression of *GCK* increased glucose uptake and promoted glucose

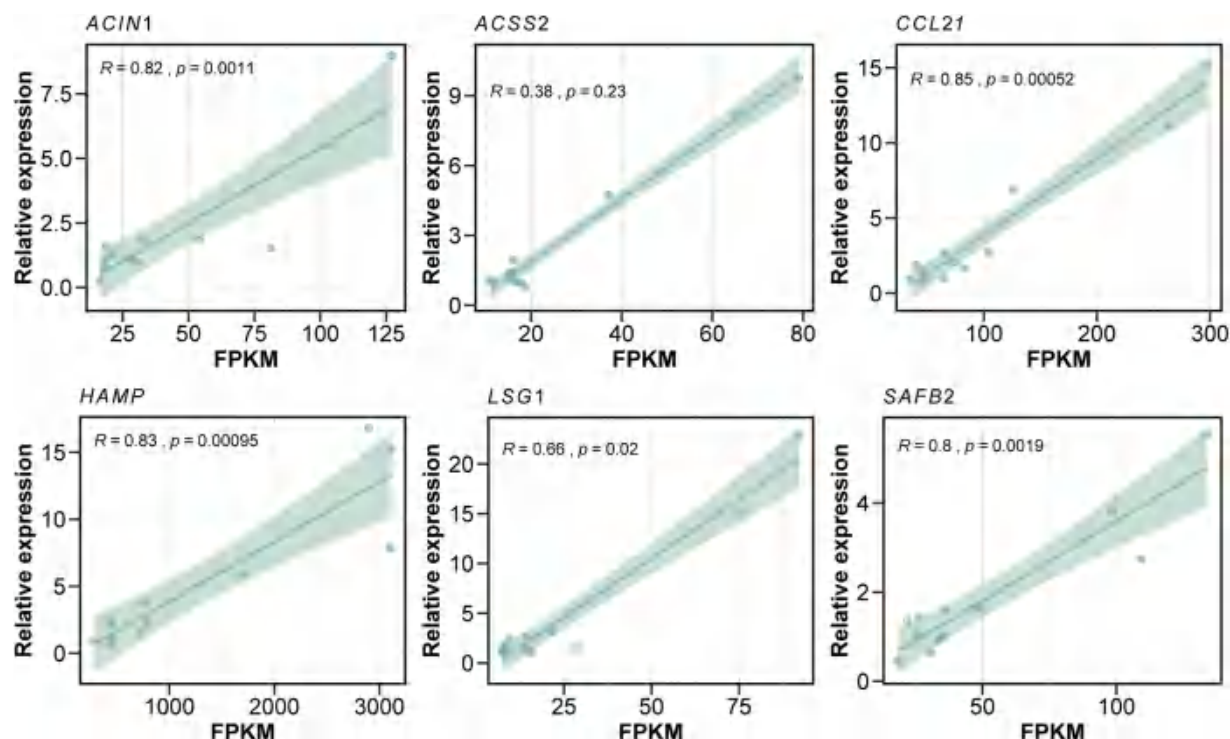


Figure 5. Correlation analysis of RNA-seq and quantitative polymerase chain reaction (qPCR) of 6 randomly selected DEGs. Six randomly selected DEGs were analyzed by real-time qPCR. The x-axis represents the fragments per kilobase of transcript per million mapped reads (FPKM) of each gene calculated by RNA-seq analysis, and the y-axis represents the relative expression of each gene calculated by qPCR. The correlation coefficient was calculated between the two measures.

utilization in the liver²⁷. Moreover, hepatic *GCK* mRNA expression was positively associated with the mRNA expression of lipogenic enzymes (*ACACB* and *FASN*) and de novo lipogenesis in the liver²⁶. Thus, the hepatic glycolytic process probably stimulated hepatic lipid synthesis. The *ENO2* gene increases glucose uptake in the liver and participates in hepatic glycolysis by converting 2-phosphoglycerate into phosphoenolpyruvate²⁸. The *HK3* (hexokinase 3) gene, one of the four hexokinase family members, is a catalytic enzyme in glycolysis²⁹. Similarly, the *PFKFB1* and *PFKFB4* genes encode key enzymes involved in glycolysis^{30,31}. In a previous study, Reyer found three adjacent SNPs in *SSC13*, two of which were located beside and in the *PFKFB4* gene in pigs, and these SNPs were significantly correlated with RFI³², suggesting that *PFKFB4* might be a candidate gene affecting FE. Liver glycolysis is one of the most important metabolic pathways regulating FE and is decreased in high-FE beef cattle³³. Moreover, a study showed that genes involved in glycolysis were downregulated in high-FE pigs³⁴.

Carbohydrate phosphorylation is the first step in glycolysis, in which glucose is phosphorylated to glucose-6-phosphate (G6P) by hexokinase³⁵. In the fed state, G6P is metabolized to generate pyruvate through glycolysis. Pyruvate is the main glycolytic product and links glycolysis to lipogenesis. Pyruvate can be completely oxidized in mitochondria to generate ATP through the citric acid cycle and oxidative phosphorylation²⁵. Glycolysis produces ATP to provide energy for growth, which is an extremely inefficient way of producing energy. In the complete oxidation of pyruvate, approximately 40% of the energy produced is used to synthesize ATP, while the remaining energy (approximately 60%) is lost as heat energy³⁶. Combining with the results of WGCNA, oxidative phosphorylation, thermogenesis, and energy metabolism-related pathways (including TCA cycle, respiratory electron transport, ATP synthesis, and heat production) were significantly enriched in MEcyan and MEpurple modules, and all of these biological processes occurred in mitochondria and are related to energy consumption^{25,37}. The MEcyan and MEpurple modules were negatively related to HE- group, indicating that higher efficiency in HE- group might due to lower energy consumption in the liver. These results implied that decreased rates of hepatic glucose metabolism, oxidative phosphorylation, thermogenesis, and energy metabolism might result in fewer energy losses in the HE- group. Hence, we hypothesized that the HE- group had more efficient energy utilization and a higher FE because the decreased glycolysis process and reduced energy losses. Correspondingly, previous analyses of FE revealed that the genes associated with the glycolytic pathway and mitochondrial activity were downregulated in the liver and muscle tissues of high-FE pigs. Less energy was lost due to the decreased glycolytic potential and mitochondrial activity in the liver, which might result in higher energy efficiency in HE- group^{19,34,38–40}.

Furthermore, pyruvate can provide a carbon source for lipogenesis in the liver. The conversion of glucose to lipids in the liver results in an approximately 23% energy loss⁴¹, which is much higher than the amount of energy lost by protein deposition in the muscle⁴². Previous studies focusing on muscle tissue revealed that high-FE pigs accumulated more muscle mass compared with low-FE pigs^{42,43}, which indicated that high-FE pigs needed to

consume more glucose in the muscle to generate ATP for protein deposition. In our study, the HE- group had lower glucose metabolic and lipid synthesizing potential in the liver, so we speculate that more glucose is consumed in the muscle tissue to provide ATP for protein deposition in the HE- group. Protein synthesis has a higher energy efficiency than lipid synthesis; thus, the HE- group exhibited higher feed efficiency. Our findings are consistent with previous studies and support the assumption that the HE- group had less heat production and greater energy utilization related to decreased glycolytic processes in the liver, which may have positive effects on FE in commercial DLY pigs.

The liver is the main organ that synthesizes fatty acids and other lipids. The precursor substances for the synthesis of fatty acids are mainly derived from glucose metabolism, especially glycolysis²⁵. In our study, lipid synthesis-related pathways were significantly enriched in the pathway enrichment analysis (among the top 10 significantly enriched pathways, 4 pathways were related to lipid synthesis), and genes involved in lipid synthesis, including *ACACB*, *CYP21A2*, *CHKA*, *SCD*, were downregulated in the HE- group. The *ACACB* gene is a known candidate gene for lipid metabolism and is related to the *de novo* synthesis of fatty acids and other lipids⁴⁴. Both *CYP21A2* and *CHKA* are involved in the synthesis of cholesterol, steroids and other lipids and are candidate genes that affect lipid synthesis in pigs^{45,46}. Lipid synthesis in the liver was negatively related to the FE trait in pigs, which has been reported in previous studies^{47,48}. Several studies found that genes involved in lipid metabolisms, such as fatty acid, steroid, and cholesterol biosynthesis, were downregulated in the liver of low-RFI (more efficient) pig^{16,49,50} and cattle⁵¹.

The PPI analysis indicated that the genes associated with lipid synthesis (including *ACACB*, *SCD*, *FASN*, *HSD17B1*, and *CYP21A2*) were candidate genes that affected the FE in commercial DLY pigs. The *SCD* gene, which encodes stearoyl CoA desaturase and promotes the synthesis of fatty acid in pigs⁵², is a candidate gene that correlates with FE¹⁴. Previous studies revealed that the *SCD* gene was downregulated in more efficient pigs^{52–54}, and the high-FE pigs had a reduction of lipid synthesis and accumulation. The *FASN* gene is a fatty acid synthase that plays an important role in the *de novo* synthesis of fatty acids⁵⁵ and was downregulated in high-FE pigs^{52–54}. In the current study, the *SCD* and *FASN* genes were downregulated in the HE- group compared with the LE- group. Moreover, by analyzing the top 5 significantly upregulated DEGs, we found that *DLK1* was related to lipid metabolism, and *DLK1* was upregulated by a log₂ (Foldchange) of 3.81 in the HE- group. Previous studies indicated that the overexpression of *DLK1* would suppress lipid synthesis, and this gene was upregulated in more efficient pigs and cattle^{51,56}. Consistent with previous studies, our results indicated that the increased FE in the HE- group might be associated with the reduction of lipid synthesis and accumulation in the liver.

Conclusion

In this study, we investigated the liver transcriptome of 6 extremely high feed efficiency pigs (HE- group) and 6 extremely low feed efficiency pigs (LE- group). A total of 507 DEGs were found between the HE- and LE- groups. GO and pathway enrichment analyses revealed that the DEGs were mainly enriched in glycolysis and lipid synthesis. The vast majority of DEGs involved in glycolysis and lipid synthesis were downregulated in the HE- group, such as *SCD*, *ACACB*, *FASN*, *GCK*, and *ENO2*. The expression patterns of these genes suggest that the related pathways might influence feed efficiency in pigs. Moreover, the results of WGCNA indicated that oxidative phosphorylation, thermogenesis, and energy metabolism-related pathways were decreased in HE- group, which resulted in higher energy efficiency in it. Briefly, the results indicated that the HE- group may have decreased glycolytic, energy consumption and lipid synthesizing potential in the liver, thereby increasing energy efficiency. Our findings provide an understanding of the molecular mechanisms in the liver in regulating the feed efficiency of DLY pigs. The key pathway and candidate genes identified in this study are potentially useful for improving porcine feed efficiency.

Methods

Ethics statement. The experimental procedures used in this study met the guidelines of the Animal Care and Use Committee of the South China Agricultural University (SCAU) (Guangzhou, People's Republic of China), and every effort was taken to minimize animal suffering. All animal experiments in this study were approved by the Animal Care and Use Committee (ACUC) of the SCAU (approval number SCAU#0030).

Animals and tissues. In this study, a total of 225 commercial Duroc × (Landrace × Yorkshire) sows, provided by Guangdong Wen's Foodstuffs Group Co., Ltd., (Yun fu, China), were selected as the experimental animals. During the experiment, the pigs were housed in an environment-controlled shed, and feed and water were offered ad libitum. The phenotypic data of 225 sows were recorded by the Osborne FIRE pig performance test system (Osborne, KS, USA). The recording period was approximately 12 weeks, during which time the weight of the animals was measured from approximately 30 kg of body weight (BW) to 100 kg BW. Each individual had a unique electronic identification tag on its ear that could be captured by an automatic feeder. Each individual's feeding time, feeding duration, feed consumption, and body weight were recorded at each visit to the feeder. The standard A-scan and contact B-scan ultrasonography were used to measure back fat (BF) of pigs in approximately 100 kg BW. The FCR and RFI were calculated for each individual during the trial. The RFI calculation method was similar to that of Cai *et al.*⁵⁷. The RFI was estimated by the linear regression of DFI on metabolic BW at mid-test (MWT), average daily gain from 30 to 100 kg (ADG), and BF. MWT was equal to [(BW at on-test + BW at off-test)/2]^{0.75}. Then, the RFI values of all individuals were ranked, and 6 pigs with extremely high-FE and 6 pigs with extremely low-FE were selected and designated as the HE- group and LE- group, respectively. Finally, a correlation analysis was conducted to calculate the correlation between RFI and FCR to verify the selection of 12 individuals.

At the end of the experiment, all 12 individuals from the HE- and LE- groups were slaughtered, and liver tissues (the middle portion of the left lateral lobe) were collected immediately. These samples were rapidly frozen in liquid nitrogen and stored at -80°C .

RNA extraction and sequencing. Total RNA was extracted from all 12 liver tissue samples using Total Kit II (OMEGA, USA) and the procedures and standards were performed according to the manual. The quantification and quality of RNA were assessed by a NanoDrop2000 microspectrophotometer (Thermo Scientific, Wilmington, DE, USA). The concentration of the mRNA ranged from 624 to 1218 ng/ μl , and the absorbance (A260/280) of all samples was between 1.8 and 2.1. Besides, an Agilent 2100 Bioanalyzer device (Agilent Technologies, Santa Clara, CA, USA) was used to assess the integrity of the RNA. The RNA integrity value (RIN) of our samples ranged between 6.2–8.8. The cDNA library was constructed using a TruSeq RNA Library Prep Kit v2 (Illumina, San Diego, CA, USA) according to the manufacturer's instructions, where mRNA was purified and enriched from 1 μg of each of the total RNA samples and then fragmented. After quality control, the libraries were sequenced on an Illumina HiSeq 4000 platform.

Read alignment and differential expression analysis. The raw reads of each sample were discriminated based on the indexing adaptors. The FastQC⁵⁸ software was used to evaluate the quality of the reads. Then, the adapter sequences and low-quality reads (the reads that adapter contamination is greater than 5 bp, Q20 ratio does not reach 85% or containing N ratios greater than 5%) were filtered out, and the high-quality reads were available for downstream analysis. The STAR: ultrafast universal RNA-seq aligner STAR_2.3.0⁵⁹ was used to map the sequencing reads with the reference genome (*Sus scrofa* 11.1), and all the options were set as STAR default values. The HTSeq⁶⁰ software was used to generate the read count tables for further differential expression analysis.

Differential expression analysis was performed between the HE- and LE- groups by using the Gfold (V1.1.2)⁶¹ software and the methods described by Audic and Claverie⁶². The Gfold (V1.1.2) software was used to count the expression level of the reads and convert them into FPKM (Fragment Per Kilobase of exon model Per Million mapped reads). The expression fold change between the two groups was calculated by the methods described by Audic and Claverie, and the Benjamini-Hochberg (BH) method was performed to calculate q -value for multiple testing. All genes were filtered by the criteria of $|\log_2(\text{Foldchange})| > 1$ and q -value < 0.001 . Genes with $\log_2(-\text{Foldchange}) > 1$ and q -value < 0.001 were defined as upregulated DEGs, while the gene with $\log_2(\text{Foldchange}) < -1$ and q -value < 0.001 were defined as downregulated DEGs.

GO and pathway enrichment analysis of DEGs. To explore the major metabolic pathways and cell signaling pathways related to FE, Gene Ontology (GO) and pathway enrichment analysis were carried out by KOBAS 3.0 (http://kobas.cbi.pku.edu.cn/anno_iden.php)⁶³, and both the Reactome and Kyoto Encyclopedia of Genes and Genomes (KEGG) database were used for pathway enrichment analysis. The BH method was applied to adjust the P -value for multiple testing. The GO terms or pathways meeting the screening criteria (with q -value < 0.05) were the significantly enriched terms or pathways.

Protein-protein interaction network construction. A protein-protein interaction (PPI) analysis of DEGs was implemented in the Search Tool for the Retrieval of Interacting Genes (STRING) database. The intensity of interaction among the input genes was evaluated and the hub gene was determined according to the degree of relationship to other genes. Then, the interaction network diagram of these genes was plotted by the Cytoscape (3.6) software⁶⁴.

Weighted gene co-expression analysis (WGCNA). Gene co-expression network analysis was performed using the R package WGCNA⁶⁵. Detailed steps are as follows. (1) Data input and cleaning: The gene expression matrix (genes that expression with variance variation accounting for top 30%) and the phenotypic matrix (including HE, LE, RFI, and FCR) for 12 individuals were used for subsequent analysis. In the phenotypic matrix, the RFI and FCR values in the matrix came from the RFI and FCR original values of each of 12 individuals. Six individuals belonging to high-FE group were marked as 1 and the remaining 6 as 0 for the "HE" values of the phenotypic matrix. Similarly, six individuals belonging to low-FE group were marked as 1 and the remaining 6 as 0 for the "LE" values of the phenotypic matrix. Gene and individual quality control were performed with the default settings of the R package WGCNA. (2) Best soft-Threshold Confirmation: Using the pickSoftThreshold function of the R package WGCNA to analyze the expression matrix obtained in the first step, the most appropriate soft threshold value was 7. (3) Network construction and module detection: Using the blockwiseModules function to analyze the expression matrix obtained in the first step for module detection with power = 7 and mergeCutHeight = 0.2, the gene set was divided into 18 modules. (4) Relating modules to phenotypic traits and identifying important genes: Correlation analysis and significance test were performed on the phenotypic matrix obtained in the first step and the modules obtained in the third step. Modules with statistically significant (p -value < 0.05) correlations were selected for downstream analysis. (5) Functional annotation of significant module genes: gene pathway analysis and gene function annotation were analyzed using R package clusterProfiler⁶⁶ with default parameters. Biological terms were considered significant if the adjusted p -value was less than 0.05.

Real-time quantitative PCR. To validate the results of the differential expression analysis from the RNA-seq data, the relative expression levels of 6 randomly selected DEGs were detected by real-time quantitative polymerase chain reaction (qPCR) technology. A total of 6 HE- and 6 LE- pigs were selected for qPCR. RNA samples were prepared by the methods mentioned above. Reverse transcription was performed by the PrimeScriptTM RT reagent kit (Takara, Japan). Then, all qPCR reactions were performed in a QuantStudioTM 7 flex device (Invitrogen Life Technologies, Carlsbad, CA, USA) following the manufacturer's instructions and three biological replicates were used in the experiment. The parameters used in the qPCR reaction were: denatured at 95°C

for 5 min; performed 40 PCR cycles (95 °C, 10 s; 60 °C, 15 s; 72 °C, 20 s); dissolution curve (95 °C, 15 s, 55 °C, 15 s, 95 °C, 15 s). Thereafter, the comparative Ct method⁶⁷ was performed to quantify the gene expression of 6 selected genes in 12 individuals. The primer sequences of these genes were designed by the Oligo 7.0 software, and the details of the primers are displayed in Table S9.

Data availability

The raw reads have been submitted to the NCBI Sequence Read Archive database (SRA) under BioProject accession number of PRJNA578377 and SRA accession number SRR10315359 - SRR10315370.

Received: 10 December 2019; Accepted: 1 June 2020;

Published online: 18 June 2020

References

1. Teagasc. Pig Herd Performance Report 2018. Foster City, CA: Teagasc Pig Development Department. <https://www.teagasc.ie/publications/2019/national-pig-herd-performance-report-2018.php> (2018).
2. Koch, R. M., Swiger, L. A., Chambers, D. & Gregory, K. E. Efficiency of Feed Use in Beef Cattle. *Journal of Animal Science* **22**, 486–494, <https://doi.org/10.2527/jas1963.222486x> (1963).
3. Do, D. N., Strathe, A. B., Jensen, J., Mark, T. & Kadarmideen, H. N. Genetic parameters for different measures of feed efficiency and related traits in boars of three pig breeds. *Journal of Animal Science* **91**, 4069–4079, <https://doi.org/10.2527/jas.2012-6197> (2013).
4. Horodyska, J., Hamill, R. M., Varley, P. F., Reyer, H. & Wimmers, K. Genome-wide association analysis and functional annotation of positional candidate genes for feed conversion efficiency and growth rate in pigs. *Plos One* **12**, e0173482, <https://doi.org/10.1371/journal.pone.0173482> (2017).
5. Onteru, S. K. *et al.* Whole genome association studies of residual feed intake and related traits in the pig. *Plos One* **8**, e61756, <https://doi.org/10.1371/journal.pone.0061756> (2013).
6. Ding, R. *et al.* Genome-wide association analysis reveals genetic loci and candidate genes for feeding behavior and eating efficiency in Duroc boars. *Plos One* **12**, e0183244, <https://doi.org/10.1371/journal.pone.0183244> (2017).
7. Quan, J. *et al.* Genome-wide association study reveals genetic loci and candidate genes for average daily gain in Duroc pigs. *Asian-Australas Journal of Animal Sciences* **31**, <https://doi.org/10.5713/ajas.17.0356> (2017).
8. Singer, J. B. Candidate Gene Association Analysis. *Methods in Molecular Biology* **573**, 223–230, https://doi.org/10.1007/978-1-60761-247-6_13 (2009).
9. Ugrappa, N. *et al.* The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344–1349, <https://doi.org/10.1126/science.1158441> (2008).
10. T Wilhelm, B. & Landry, J.-R. RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing. *Volume* **48**, <https://doi.org/10.1016/j.ymeth.2009.03.016> (2009).
11. Alexandre, P. A. *et al.* Liver transcriptomic networks reveal main biological processes associated with feed efficiency in beef cattle. *BMC Genomics* **16**, 1073, <https://doi.org/10.1186/s12864-015-2292-8> (2015).
12. Kong, R. S. G., Liang, G., Chen, Y., Stothard, P. & Le, L. G. Transcriptome profiling of the rumen epithelium of beef cattle differing in residual feed intake. *BMC Genomics* **17**, 592, <https://doi.org/10.1186/s12864-016-2935-4> (2016).
13. Zhou, N. Using RNA-seq to characterize the biological basis of variation in feed efficiency in broiler chickens, <http://udspace.udel.edu/handle/19716/17035> (2015).
14. Bartz, M. *et al.* Transcript abundance of the pig stearoyl-CoA desaturase gene has no effect on fatty acid composition in muscle and fat tissues, but its polymorphism within the putative microRNA target site is associated with daily body weight gain and feed conversion ratio. *Journal of Animal Science* **91**, 10–19, <https://doi.org/10.2527/jas.2012-5380> (2013).
15. Zhao, Y. *et al.* Transcriptome analysis reveals that Vitamin A metabolism in the liver affects feed efficiency in pigs. *G3 GenesGenetics* **6**, 3615–3624, <https://doi.org/10.1534/g3.116.032839> (2016).
16. Ramayocaldas, Y. *et al.* Integrative approach using liver and duodenum RNA-Seq data identifies candidate genes and pathways associated with feed efficiency in pigs. *Scientific Reports* **8**, <https://doi.org/10.1038/s41598-017-19072-5> (2018).
17. Dentin, R., Girard, J. & Postic, C. Carbohydrate responsive element binding protein (ChREBP) and sterol regulatory element binding protein-1c (SREBP-1c): two key regulators of glucose metabolism and lipid synthesis in liver. *Biochimie* **87**, 81–86, <https://doi.org/10.1016/j.biochi.2004.11.008> (2005).
18. Kosaku, U. & Repa, J. J. Carbohydrate response element binding protein, ChREBP, a transcription factor coupling hepatic glucose utilization and lipid synthesis. *Cell Metabolism* **4**, 107–110, <https://doi.org/10.1016/j.cmet.2006.06.008> (2006).
19. Reyer, H. *et al.* Strategies towards improved feed efficiency in pigs comprise molecular shifts in hepatic lipid and carbohydrate metabolism. *International Journal of Molecular Sciences* **18**, 1674, <https://doi.org/10.3390/ijms18081674> (2017).
20. Madeira, M. S. *et al.* Restriction of dietary protein does not promote hepatic lipogenesis in lean or fatty pigs. *British Journal of Nutrition* **115**, 1, <https://doi.org/10.1017/S0007114516000453> (2016).
21. Lkhagvadorj, S. *et al.* Gene expression profiling of the short-term adaptive response to acute caloric restriction in liver and adipose tissues of pigs differing in feed efficiency. *American Journal of Physiology Regulatory Integrative and Comparative Physiology* **298**, R494, <https://doi.org/10.1152/ajpregu.00632.2009> (2010).
22. Horodyska, J. *et al.* RNA-Seq of liver from pigs divergent in feed efficiency highlights shifts in macronutrient metabolism, hepatic growth and immune response. *Frontiers in Genetics* **10**, <https://doi.org/10.3389/fgene.2019.00117> (2019).
23. Quan, J. *et al.* Exploring the fecal microbial composition and metagenomic functional capacities associated with feed efficiency in commercial DLY Pigs. *Frontiers in Microbiology* **10**, <https://doi.org/10.3389/fmicb.2019.00052> (2019).
24. Heng, Q. *et al.* Isolation of the novel human guanine nucleotide exchange factor Src homology 3 domain-containing guanine nucleotide exchange factor (SGEF) and of C-terminal SGEF, an N-terminally truncated form of SGEF, the expression of which is regulated by androgen in prostate cancer cells. *Endocrinology* **144**, 1742–1752, <https://doi.org/10.1210/en.2002-220984> (2003).
25. Rui, L. Energy metabolism in the liver. *Comprehensive Physiology* **4**, 177–197, <https://doi.org/10.1002/cphy.c130024> (2014).
26. Peter, A. *et al.* Hepatic glucokinase expression is associated with lipogenesis and fatty liver in humans. *Journal Of Clinical Endocrinology & Metabolism* **96**, E1126, <https://doi.org/10.1210/jc.2010-2017> (2011).
27. Torres, T. P. *et al.* Restoration of hepatic glucokinase expression corrects hepatic glucose flux and normalizes plasma glucose in Zucker diabetic fatty rats. *Diabetes* **58**, 78–86, <https://doi.org/10.2337/db08-1119> (2009).
28. Baurhoo, B. Reduction of Salmonella-induced enteric and systemic inflammation by mannan-oligosaccharide prebiotic through improvement of innate defense mechanism. (McGill University (Canada). (2012).
29. Nakamura, N., Shibata, H., O'Brien, D. A., Mori, C. & Eddy, E. M. Spermatogenic cell-specific type 1 hexokinase is the predominant hexokinase in sperm. *Molecular Reproduction & Development* **75**, 632–640, <https://doi.org/10.1002/mrd.20791> (2010).
30. Hao, Y., Cui, Y. & Gu, X. Genome-wide DNA methylation profiles changes associated with constant heat stress in pigs as measured by bisulfite sequencing. *Scientific Reports* **6**, 27507, <https://doi.org/10.1038/srep27507> (2016).
31. Yun, S. J. *et al.* PFKFB4 as a prognostic marker in non-muscle-invasive bladder cancer. *Urologic Oncology* **30**, 893–899, <https://doi.org/10.1016/j.urolonc.2010.08.018> (2012).

32. Reyer, H. *et al.* Exploring the genetics of feed efficiency and feeding behaviour traits in a pig line highly selected for performance characteristics. *Molecular Genetics & Genomics* **292**, 1001–1011, <https://doi.org/10.1007/s00438-017-1325-1> (2017).
33. Fonseca, L. D. *et al.* Liver proteomics unravel the metabolic pathways related to feed efficiency in beef cattle. *Scientific Reports* **9**, 5364, <https://doi.org/10.1038/s41598-019-41813-x> (2019).
34. Jing, L. *et al.* Transcriptome analysis of mRNA and miRNA in skeletal muscle indicates an important network for differential residual feed intake in pigs. *Scientific Reports* **5**, 11953, <https://doi.org/10.1038/srep11953> (2015).
35. Faizeh, A. Q. & Mathias, M. CK2 and the regulation of the carbohydrate metabolism. *Metabolism-clinical and Experimental* **61**, 1512–1517, <https://doi.org/10.1016/j.metabol.2012.07.011> (2012).
36. Baldwin, R. L. & Donovan, K. C. Modeling Ruminant Digestion and Metabolism. *Advances in Experimental Medicine & Biology* **445**, 325, https://doi.org/10.1007/978-1-4899-1959-5_21 (1998).
37. Busiello, R. A., Savarese, S. & Lombardi, A. Mitochondrial uncoupling proteins and energy metabolism. *Frontiers in Physiology* **6**, 36, <https://doi.org/10.3389/fphys.2015.00036> (2015).
38. Boddicker, N., Gabler, N. K., Spurlock, M. E., Nettleton, D. & Dekkers, J. C. M. Effects of ad libitum and restricted feeding on early production performance and body composition of Yorkshire pigs selected for reduced residual feed intake. *Animal* **5**, 1344–1353, <https://doi.org/10.1017/S17517311100036X> (2011).
39. Barea, R. *et al.* Energy utilization in pigs selected for high and low residual feed intake. *Journal of Animal Science* **88**, 2062–2072, <https://doi.org/10.2527/jas.2009-2395> (2010).
40. Fu, L. *et al.* Proteomic analysis indicates that mitochondrial energy metabolism in skeletal muscle tissue is negatively correlated with feed efficiency in pigs. *Scientific Reports* **7**, 45291, <https://doi.org/10.1038/srep45291> (2017).
41. Patience, J. *The metabolic basis of feed-energy efficiency in swine*, <https://conservancy.umn.edu/bitstream/handle/11299/139290/Gutierrez.pdf?sequence=1> (2012).
42. Faure, J. *et al.* Consequences of divergent selection for residual feed intake in pigs on muscle energy metabolism and meat quality. *Meat Science* **93**, 37–45, <https://doi.org/10.1016/j.meatsci.2012.07.006> (2013).
43. Vincent, A. *et al.* Divergent selection for residual feed intake affects the transcriptomic and proteomic profiles of pig skeletal muscle. *Journal of Animal Science* **93**, 2745–2758, <https://doi.org/10.2527/jas.2015-8928> (2015).
44. Lixin, Z. *et al.* Lipid in the livers of adolescents with nonalcoholic steatohepatitis: combined effects of pathways on steatosis. *Metabolism Clinical & Experimental* **60**, 1001–1011, <https://doi.org/10.1016/j.metabol.2010.10.003> (2011).
45. Xing, K. *et al.* Identification of genes for controlling swine adipose deposition by integrating transcriptome, whole-genome resequencing, and quantitative trait loci data. *Scientific Reports* **6**, 23219, <https://doi.org/10.1038/srep23219> (2016).
46. Fu, Y. *et al.* Integrated analysis of methylome, transcriptome and miRNAome of three pig breeds. *Epigenomics* **10**, epi-2017–0087, <https://doi.org/10.2217/epi-2017-0087> (2018).
47. Liu, P. *et al.* Influence of thermally oxidized vegetable oils and animal fats on growth performance, liver gene expression, and liver and serum cholesterol and triglycerides in young pigs. *Journal of Animal Science* **92**, 2960–2970, <https://doi.org/10.2527/jas.2012-5709> (2014).
48. Madeira, M. S. *et al.* Restriction of dietary protein does not promote hepatic lipogenesis in lean or fatty pigs. *British Journal of Nutrition* **115**, 1, <https://doi.org/10.1017/S0007114516000453> (2016).
49. Lkhagvadorj, S. *et al.* Gene expression profiling of the short-term adaptive response to acute caloric restriction in liver and adipose tissues of pigs differing in feed efficiency. *American Journal of Physiology Regulatory Integrative & Comparative Physiology* **298**, R494, <https://doi.org/10.1152/ajpregu.00632.2009> (2010).
50. Gondret, F. *et al.* A transcriptome multi-tissue analysis identifies biological pathways and genes associated with variations in feed efficiency of growing pigs. *BMC Genomics* **18**, 244, <https://doi.org/10.1186/s12864-017-3639-0> (2017).
51. Mukibi, R. *et al.* Transcriptome analyses reveal reduced hepatic lipid synthesis and accumulation in more feed efficient beef cattle. *Scientific Reports* **8**, 7303, <https://doi.org/10.1038/s41598-018-25605-3> (2018).
52. Wang, Z. *et al.* Correction: Identification of genes related to growth and lipid deposition from transcriptome profiles of pig muscle tissue. *Plos One* **12**, e0172930, <https://doi.org/10.1371/journal.pone.0172930> (2017).
53. Ghosh, M. *et al.* Evaluation of body growth and immunity-related differentially expressed genes through deep RNA sequencing in the piglets of Jeju native pig and Berkshire. *Animal Genetics* **46**, 255–264, <https://doi.org/10.1111/age.12281> (2015).
54. Blankenship, K., Gilley, A., Piekarski, A., Orlowski, S. & Dridi, S. Differential expression of feeding-related hypothalamic neuropeptides in the first generation of quails divergently selected for low or high feed efficiency. *Neuropeptides* **58**, 31–40, <https://doi.org/10.1016/j.npep.2015.12.007> (2016).
55. Towle, H. C., Kaytor, E. N. & Shih, H. M. Regulation of the expression of lipogenic enzyme genes by carbohydrate. *Annual Review of Nutrition* **17**, 405–433, <https://doi.org/10.1146/annurev.nutr.17.1.405> (2003).
56. Kim, K. S., Kim, J. J., Dekkers, J. C. M. & Rothschild, M. F. Polar overdominant inheritance of a DLK1 polymorphism is associated with growth and fatness in pigs. *Mammalian Genome* **15**, 552–559, <https://doi.org/10.1007/s00335-004-2341-0> (2004).
57. Cai, W., Casey, D. S. & Dekkers, J. C. M. Selection response and genetic parameters for residual feed intake in Yorkshire swine. *Journal of Animal Science* **86**, 287, <https://doi.org/10.2527/jas.2007-0396> (2008).
58. Andrews, S. FastQC: a quality control tool for high throughput sequence data, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (2010).
59. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21, <https://doi.org/10.1093/bioinformatics/bts635> (2013).
60. Anders, S., Theodor, P. & Huber, W. H. HTseq: a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169, <https://doi.org/10.1093/bioinformatics/btu638> (2014).
61. Feng, J. *et al.* GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. *Bioinformatics* **28**, 2782–2788, <https://doi.org/10.1093/bioinformatics/bts515> (2012).
62. Audic, S. & Claverie, J. M. The significance of digital gene expression profiles. *Genome Research* **7**, 986–995, <https://doi.org/10.1101/gr.7.10.986> (1997).
63. Chen, X. *et al.* KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Research* **39**, 316–322, <https://doi.org/10.1093/nar/gkr483> (2011).
64. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **13**, 2498–2504, <https://doi.org/10.1101/gr.1239303> (2003).
65. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559, <https://doi.org/10.1186/1471-2105-9-559> (2008).
66. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology* **16**, 284–287, <https://doi.org/10.1089/omi.2011.0118> (2012).
67. Thomas, S., Kenneth, D. & Livak, J. Analyzing real-time PCR data by comparative CT method. *Nature Protocols* **3**, <https://doi.org/10.1038/nprot.2008.73> (2008).

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant No. 31972540), the Project of Swine Innovation Team in Guangdong Modern Agricultural Research System (Grant No. 2019KJ26), the Pearl River Nova Program of Guangzhou (Grant No. 201906010011) and the Natural Science Foundation of Guangdong Province (Grant No. 2017A030313213).

Author contributions

Z.W., E.Z. and J.Y. conceived and designed the experiments. C.X., X.W., Z.Z., J.W., S.Z., J.Q., R.D., Y.Y., L.P., Z.W. and E.Z. collected the samples and recorded the phenotypes. C.X. and J.W. performed the P.C.R. C.X. and X.W. analyzed the data. C.X., X.W. and J.Y. wrote and revised the manuscript. Z.W., E.Z. and J.Y. contributed the materials. All authors reviewed and approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-66988-6>.

Correspondence and requests for materials should be addressed to E.Z. or J.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

Brain Transcriptome Analysis Reveals Potential Transcription Factors and Biological Pathways Associated with Feed Efficiency in Commercial DLY Pigs

Cineng Xu,* Xingwang Wang,* Shenping Zhou, Jie Wu, Qian Geng, Donglin Ruan, Yibin Qiu, Jianping Quan, Rongrong Ding, Gengyuan Cai, Zhenfang Wu, Enqin Zheng, and Jie Yang

Feed efficiency (FE) is one of the most important economic traits in the porcine industry. In this study, high-throughput RNA sequencing (RNA-seq) was first utilized for brain tissue transcriptome analysis in pigs to indicate the potential genes and biological pathways related to FE in pigs. A total of 8 pigs with either extremely high-FE group (HE-group) or low-FE group (LE-group) were selected from 225 Duroc × (Landrace × Yorkshire) (DLY) pigs for transcriptomic analysis. RNA-seq analysis was performed to determine differentially expressed genes (DEGs) between the HE- and LE-group, and 430 DEGs were identified in brain tissues of pigs ($|\log_2(\text{FoldChange})| > 1$; adjusted p -values < 0.05). Gene Ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis showed that the DEGs were mainly enriched in synaptic signaling or transmission, and hormone secretion pathways, in which insulin secretion, and oxytocin signaling pathways were closely associated with FE by regulating feeding behavior and energy metabolism (adjusted p -values < 0.05). Further, the transcription factors (TFs) analysis and gene co-expression network analysis indicated three hub differentially expressed TFs (NR2F2, TFAP2D, and HNF1B) that affected FE by mainly regulating feeding behavior, insulin sensitivity, or energy metabolism. Our findings suggest several potential TFs and biological pathways for further investigations of FE in pigs.

Keywords: transcriptome analysis, DLY pig, feed efficiency, brain tissue, transcription factor

Introduction

Feed accounts for about 70% of the total production cost in pig production (Teagasc, 2018). Improving feed efficiency (FE) is a major consideration for porcine industry, as it will simultaneously reduce feed costs and environmental impact.

Residual feed intake (RFI) has been considered the most suitable measure of FE, which is defined as the difference between actual feed intake and the expected feed requirements for maintenance of body metabolic processes and production (Koch et al., 1963). Estimates of the heritability of RFI range from 0.14 to 0.53 (Gilbert et al., 2007; Cai et al., 2008). Previous studies showed that low RFI pigs had less average daily feed intake (DFI) and exhibited an increase in FE (Cai et al., 2008; Boddicker et al., 2011). Selecting for low RFI individuals can ef-

fectively improve FE in pig while causing no correlated change in other economic traits such as carcass lean ratio and eye muscle area (Hoque and Suzuki, 2009; Patience et al., 2015).

To unravel the candidate genes and molecular mechanisms of FE, several transcriptome profiling methods by RNA sequencing (RNA-seq) have been performed in several tissues, including the liver, skeletal muscle, and small intestine (Horodyska et al., 2018; Wang et al., 2019; Xu et al., 2020). Previous studies indicated that FE-divergent pigs shifted in biological processes (BPs), including energy metabolism, carbohydrate and lipid metabolism, nutrient digestion and absorption, and immune responses (Rakhshandeh et al., 2012; Jing et al., 2015; Gondret et al., 2017). Although these studies have provided insights on the molecular mechanisms that underlie FE in pigs, the results pertaining to them are not always consistent and the regulation of FE is still

College of Animal Science and National Engineering Research Center for Breeding Swine Industry, South China Agricultural University, Guangdong, P.R. China.

*These authors contributed equally to this work.

ambiguous. In this regard, further study is needed to elucidate the functional mechanisms that govern FE.

To date, a transcriptome study has been conducted on the brain of pigs to identify allele-specific expression (Oczkowicz et al., 2018), but none of the brain transcriptome studies has been performed to unravel the candidate genes and molecular mechanism of FE in pigs, whereas the brain plays a crucial role in regulating feed intake and energy metabolism. The brain central nervous system (CNS) integrates with other organs or tissues, including liver, gastrointestinal tract, pancreas, and adipose tissue, to form the circuits to regulate appetite and energy balance (Mcminn et al., 2000; Berthoud, 2002; Blevins et al., 2002). The brain receives peptides, hormones, enzymes, receptors, and bioactive factors from peripheral tissues and thereafter generates appropriate responses (Woods et al., 1998; Mcminn et al., 2000).

For example, insulin produced by the pancreas binds to insulin-binding sites localized to nuclei of the brain, which are related to energy homeostasis, then suppresses food intake, and finally enhances peripheral glucose metabolism (Baskin et al., 1986; Obici et al., 2002; Obici and Rossetti, 2003). Adipose tissue secretes leptin, which binds to the receptors located on key neurons after it enters the CNS and thus the expression of neuropeptides that regulate energy balance is altered (Mcminn et al., 2000; Trayhurn et al., 2006). In the stages of the digestion, absorption, and metabolism of nutrients, specific sensors in the digestive tract are able to carry signals to the brain. In addition, the gastrointestinal tract communicates intensely with the brain through the hormones, such as CCK and ghrelin, to regulate food intake and energy metabolism (Berthoud, 2002).

Based on the important role of the brain in regulating energy metabolism and food intake, we performed a brain transcriptome analysis of high-FE group (HE-group) and low-FE group (LE-group) to identify candidate genes and pathways that significantly correlated with the FE of pigs. Further, the identified genes and pathways that affect FE can provide theoretical support for pig selection, to improve the FE and economic benefits in commercial pig production in the future.

Materials and Methods

Ethics statement

The experimental procedures of this study met the guidelines of the Animal Care and Use Committee of the South China Agricultural University (ACUC of the SCAU) (Guangzhou, People's Republic of China). All animal experiments in this study were approved by ACUC of the SCAU (approval number SCAU#0030), and every effort was taken to minimize animal suffering.

Animals

The experiment included 225 DLY sows provided by Guangdong Wen's Foodstuffs Group Co., Ltd., Yun fu, China. The Osborne FIRE pig performance test system (Osborne, KS) were used to continuously record the phenotypic data of pigs from 30 kg of body weight (BW) to 100 kg BW (approximately 12 weeks). All of them had a unique electronic identification tag on the ear that could be captured by an automatic feeder and their feeding time,

feeding duration, feed intake, and BW were recorded at each visit to the feeder. Throughout the experimental period, all pigs were housed in an environment-controlled shed, and feed and water were offered ad libitum.

The RFI and feed conversion ratio (FCR) of all individuals were calculated after the experiment. The RFI value was estimated by the linear regression of DFI on metabolic BW at mid-test (MWT), average daily gain (ADG) from 30 to 100 kg, and back fat. MWT was equal to $[(BW \text{ at on-test} + BW \text{ at off-test})/2]^{0.75}$ (Cai et al., 2008). FCR calculated as DFI divided by ADG during a specified period. Thereafter, according to the RFI values of all 225 individuals, 4 extremely low RFI pigs and 4 extremely high RFI pigs were assigned as HE- and LE-groups.

At the end of the whole experiment period, all selected eight individuals from the HE- and LE-groups were slaughtered, and the cerebral cortex of brain tissues was sampled immediately. These samples were rapidly frozen in liquid nitrogen and stored at -80°C until further processing.

RNA Extraction and RNA-seq

Total RNA was extracted from all eight brain tissue samples by using RNAout 1.0 (Tianenze, Beijing, China) by Novogene, Beijing, China. The concentration and purity of total RNA were assessed by a NanoDrop spectrophotometer (Thermo Scientific, Wilmington, DE). The concentration of the mRNA ranged from 56 to 370 ng/mL. The RNA integrity value of all mRNA ranged from 6.3 to 8.6, which was assessed by an Agilent 2100 Bioanalyzer device (Agilent Technologies, Santa Clara, CA). Library preparation and RNA-Seq were conducted at Beijing Novogene Biological Information Technology Co., Ltd., Beijing, China (www.novogene.com). The paired-end libraries were prepared by Illumina TruSeq™ RNA Library Preparation Kit v2 (Illumina, San Diego) and were sequenced following the standard procedure on an Illumina HiSeq platform, and paired-end reads were generated.

RNA-seq data analysis

The raw reads were quality controlled before mapping to the reference genome. The FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) program was used to evaluate the quality of the raw reads. Then, the raw reads were filtered to remove adapter sequences and the reads wherein adapter contamination is greater than 5 bp, Q20 ratio does not reach 85% or contain N ratios greater than 5%. The high-quality reads were mapped onto the reference genome (*Sus scrofa* 11.1) by using the STAR (Alexander et al., 2013): ultrafast universal RNA-seq aligner STAR 2.3.059 with default parameters. Finally, the read count tables were generated by the HTSeq (Simon et al., 2015) software and the fragment per kilobases of transcript per million reads (FPKM) of each gene was calculated based on the length of the gene and read count tables.

The Bioconductor package DESeq2 (Love et al., 2014) was performed to identify differentially expressed genes (DEGs) in the brain transcriptome between HE- and LE-groups. Raw gene count data were provided to DESeq2 (Supplementary Table S1), which normalized the data and calculated the expression fold change between the two groups, and the Benjamini and Hochberg method was used to adjust the p-value for multiple tests. The DEGs were defined as those genes with a $|\log_2(\text{FoldChange})| > 1$ and an adjusted p-value < 0.05 .

Functional enrichment analysis and pathways analysis

Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment analysis of DEGs were implemented by the clusterProfiler (3.14) R package (Bioconductor) (Yu et al., 2012). The ENTREZ gene IDs were inputted, and genome-wide annotation for pig was obtained by the “org.Ss.eg.db” package. The GO enrichment analysis was performed by the enrichGO function, in which DEGs were divided into three groups: molecular function (MF), cellular component (CC), and BP. The enrichKEGG function was used for KEGG pathways enrichment analysis. An adjusted p-value <0.05 was the threshold for significantly enriched GO terms and KEGG pathways.

Identification of key transcription factors

Transcription factors (TFs) can potentially affect the DEGs in the dataset, which can be identified by regulatory impact factors (RIF) metrics described by Reverter et al. (2010). To identify TFs in all DEGs, all the 1490 TFs in 71 families of *Sus scrofa* were obtained at AnimalTFDB 3.0 (<http://bioinfo.life.hust.edu.cn/AnimalTFDB/>) (Hu et al., 2019). Then, the read count data were converted to transcripts per million reads (TPM) and the mean values for each gene were used as a threshold to filter the DEGs. The DEGs with values above half of the previous means were used for RIF and partial correlation with information theory (PCIT) analyses.

The RIF analysis was performed by the RIF function of the CeTF R package (Bioconductor) (De Biagi et al., 2020), which outputted two different and inter-connected RIF scores (RIF1 and RIF2). The RIF1 metric classifies the TFs as most differentially co-expressed with the highly abundant and highly DE genes, whereas the RIF2 metric classifies the TF showing evidence as predictors of change in abundance of DEGs between two groups. A TF was deemed as a key TF if $|RIF1|$ or $|RIF2| > 1.96$ (Reverter et al., 2010).

Gene co-expression network

Significant pair-wise gene interactions were identified with PCIT analysis to construct a co-expression gene network. The PCIT function from the CeTF R package (De Biagi et al., 2020) was used for PCIT analysis, and the normalized expression data (TPM) as previously described was used as input. Then, the pairwise correlations of all differentially expressed TFs and their targets (DEGs) were retained and a significant partial correlation $|r| > 0.95$ was selected to construct gene co-expression networks. Finally, the gene co-expression network was visualized on the Cytoscape software (3.6.0) (Shannon et al., 2003).

Real-time quantitative polymerase chain reaction validation

To validate the reliability and accuracy of RNA-seq data, a total of six DEGs were randomly selected for real-time quantitative polymerase chain reaction (RT-qPCR). The primers of these genes were designed by the Oligo 7.0 software, and encoding b-actin was used as the reference gene (Supplementary Table S2). Then, RT-qPCR analysis was performed in

a QuantStudio™ 7 flex device (Invitrogen Life Technologies, Carlsbad, CA) following the manufacturer's instructions, and each sample was run in triplicate for the experiment. The parameters were: denaturation at 95°C for 5 min, then performance of 40 PCR cycles (95°C, 10 s; 60°C, 15 s; 72°C, 20 s). Dissolution curve analysis was performed (95°C, 15 s, 55°C, 15 s, 95°C, 15 s) to validate the PCR specificity after the qRT-PCR cycles. Thereafter, the comparative Ct method (Schmittgen and Livak, 2008) was used to quantify the relative expression levels of six DEGs normalized to b-actin.

Data availability

The raw reads of eight samples have been submitted to the NCBI Sequence Read Archive database (SRA) with the BioProject accession number of PRJNA645684 and SRA accession number SRR12202382–SRR12202389.

Result

Animals

Differences in RFI and other phenotypic traits are shown in Table 1. Two extreme groups (n=8) were selected with low or high RFI. The HE-group had RFI values of -0.2 ± 0.07 kg/day, compared with the RFI value of 0.17 ± 0.09 kg/day for the LE-group. The two groups differed significantly in FCR as well; the FCR value was lower in the HE-group, which had a positive correlation between RFI ($r=0.99$; p-value = 3.19×10^{-6}). Importantly, ADG and average metabolic BW gain did not differ between the HE-group and the LE-group (p-value = 7.25×10^{-1}). However, DFI in the HE-group was significantly lower than in the LE-group (p-value = 6.96×10^{-3}), which indicated that the lower RFI in the HE-group was achieved by a reduced DFI without compromising growth (Table 1 and Fig. 1).

Mapping statistics summary

A total of 30.4 Gb raw data were generated from eight samples after sequencing using the Illumina HiSeq-PE150 platform. After filtering low-quality reads, the number of

Table 1. Feed Efficiency Phenotypic Parameters in Duroc × (Landrace × Yorkshire) Pigs from High-Feed Efficiency Group and Low-Feed Efficiency Group

Trait	HE-group (n=4)	LE-group (n=4)	p-Value
SW (kg)	30.25 ± 0.70	29.75 ± 0.25	2.57E-01
EW (kg)	99.92 ± 0.53	99.63 ± 1.07	6.40E-01
TD (day)	86.00 ± 2.45	88.75 ± 3.30	2.34E-01
ADG (kg/day)	0.81 ± 0.02	0.79 ± 0.03	4.22E-01
AMBW (kg)	24.12 ± 0.14	24.17 ± 0.24	7.25E-01
DFI (kg)	1.76 ± 0.05	2.09 ± 0.12	6.96E-03
FCR (kg/kg)	2.19 ± 0.10	2.68 ± 0.07	3.83E-04
RFI (kg/day)	-0.20 ± 0.07	0.17 ± 0.09	7.74E-04

ADG, average daily gain; AMBW, average metabolic body weight gain; DFI, daily feed intake; EW, ending weight; FCR, feed conversion ratio; FE, feed efficiency; HE-Group, high-FE group; LE-Group, low-FE group; RFI, residual feed intake; SW, starting weight; TD, testing days.

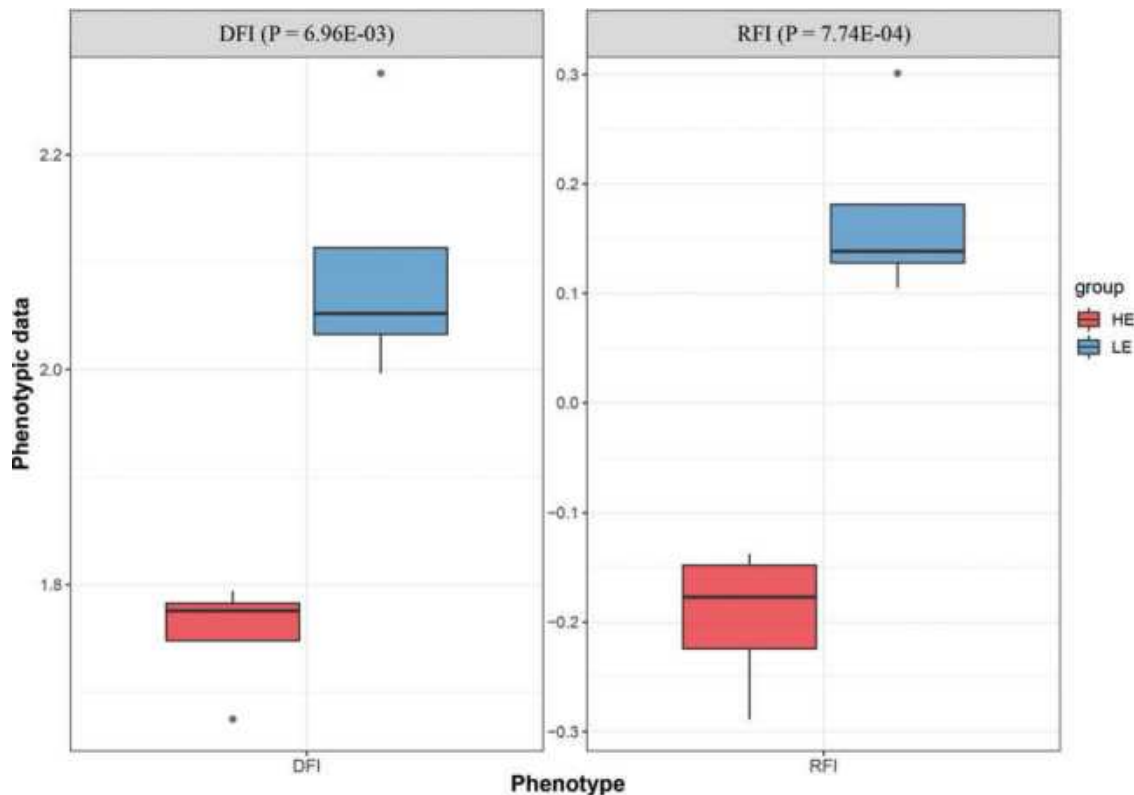


FIG. 1. Boxplot of the DFI and RFI in pigs from the HE-group and the LE-group. Red and blue color in the boxplot represent the HE-group and the LE-group, respectively. Both DFI and RFI in the HE-group were significantly lower than the LE-group. FE phenotypic parameters in DLY pigs from the HE-group and the LE-group. DFI, daily feed intake; DLY, Duroc \times (Landrace \times Yorkshire); FE, feed efficiency; HE-group, high FE group; LE-group, low FE group; RFI, residual feed intake. Color images are available online.

clean reads for eight samples ranged from 23,247,269 to 29,173,265. The percentage of reads mapped to the reference genome ranged from 96.51% to 97.83%, and the percentage of unique mapped reads ranged from 93.94% to 95.45% (Supplementary Table S3).

A total of 430 DEGs were identified between the HE- and the LE-group

In this study, adjusted p -value < 0.05 & $|\log_2(\text{FoldChange})| > 1$ were used as the criteria for significant differences in gene expression. Overall, a total of 430 DEGs were identified between two groups, with 229 upregulated and 201 downregulated genes in the HE-group (Fig. 2). All differentially expressed TFs were highlighted in the volcano plots, including TFAP2D, SCRT1, RORB, RARB, POU6F2, NR2F2, NEUROD1, MEF2A, IRX1, HNF1B, HES3, FOSB, ETV4, ESRRG, EMX1, DRGX, CEBPD, and BARX2. The details of all DEGs with their $\log_2(\text{Foldchange})$ and adjusted p -value are listed in Supplementary Table S4.

GO analysis of DEGs

GO analysis was performed to further explore the signaling pathways that the DEGs participate in. According to the result of GO analysis, the DEGs were distributed across 100 significantly enriched functional terms (adjusted

p -value < 0.05), including 28 BP terms, 39 MF terms, and 33 CC terms. All significantly enriched pathways and the genes contained in each pathway are listed in Supplementary Table S5. The top 10 significantly enriched BP, MF, and CC terms are shown in the barplot (Fig. 3). The most significantly enriched term was the G protein-coupled receptor signaling pathway (GO:0007186), which includes peptide and nonpeptide hormones and neurotransmitters, proteinases, lipids, and growth factors. Among the 10 significantly enriched GO_BP terms, seven terms were associated with synaptic signaling or transmission.

KEGG enrichment analysis of DEGs

To explore the major metabolic pathways related to FE, KEGG enrichment analysis was carried out, in which all DEGs were mapped to a reference pathway in the KEGG database to determine the biological pathways that these genes may be involved in. The DEGs were significantly enriched in 18 pathways (adjusted p -value < 0.05), which are shown in the KEGG scatter plot (Fig. 4), and the details of the results are listed in Supplementary Table S6. The results revealed that the calcium signaling pathway and neuroactive ligand-receptor interaction are the most significantly enriched pathways. Moreover, those 18 pathways were mainly related to hormone secretion, including insulin secretion, GnRH secretion, aldosterone synthesis and secretion, Oxytocin signaling pathway,

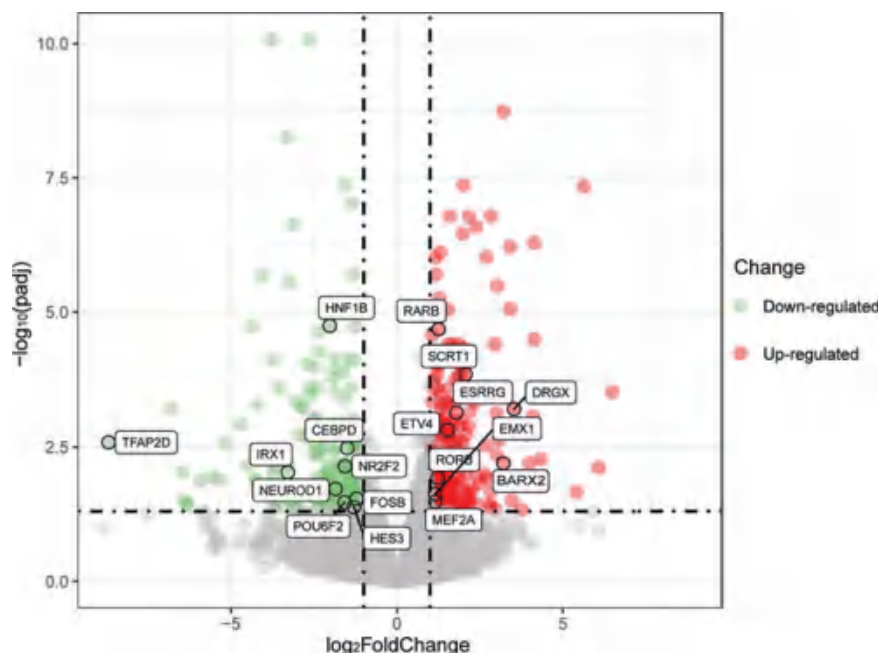


FIG. 2. DEGs between HE- and LE-groups. A plot of DEGs with $|\log_2(\text{Foldchange})| > 1$ and adjusted p-value < 0.05 . The x-axis and y-axis represent the $\log_2(\text{Foldchange})$ and $-\log_{10}(\text{adjusted p-value})$, respectively. Red dots represent up-regulated DEGs, and green dots represent down-regulated DEGs. The genes marked in the figure are 18 differentially expressed TFs between two groups. DEGs, differentially expressed genes; TFs, transcription factors. Color images are available online.

and pancreatic secretion, and almost all genes involved in these pathways were upregulated, which might improve the FE of pigs.

TFs analysis

In the current study, the prediction of TFs with RIF analysis for the whole dataset of DEGs was performed and 18 differentially expressed TFs were found among 430

DEGs, which showed differences in expression level between HE- and LE-groups. The differentially expressed TFs included members of the AP-2, RXR-like, zf-C2H2, THR-like, Pou, bHLH, SRF, Homeobox, TF-bZIP, ETS, and ESR-like gene families. The RIF1 and RIF2 scores of 18 differentially expressed TFs are shown in Supplementary Table S7. A total of 5 TFs had negative values of RIF1 score, whereas 13 TFs had positive values. For RIF2 scores, a total of 11 TFs had negative values of RIF2 score, and 7

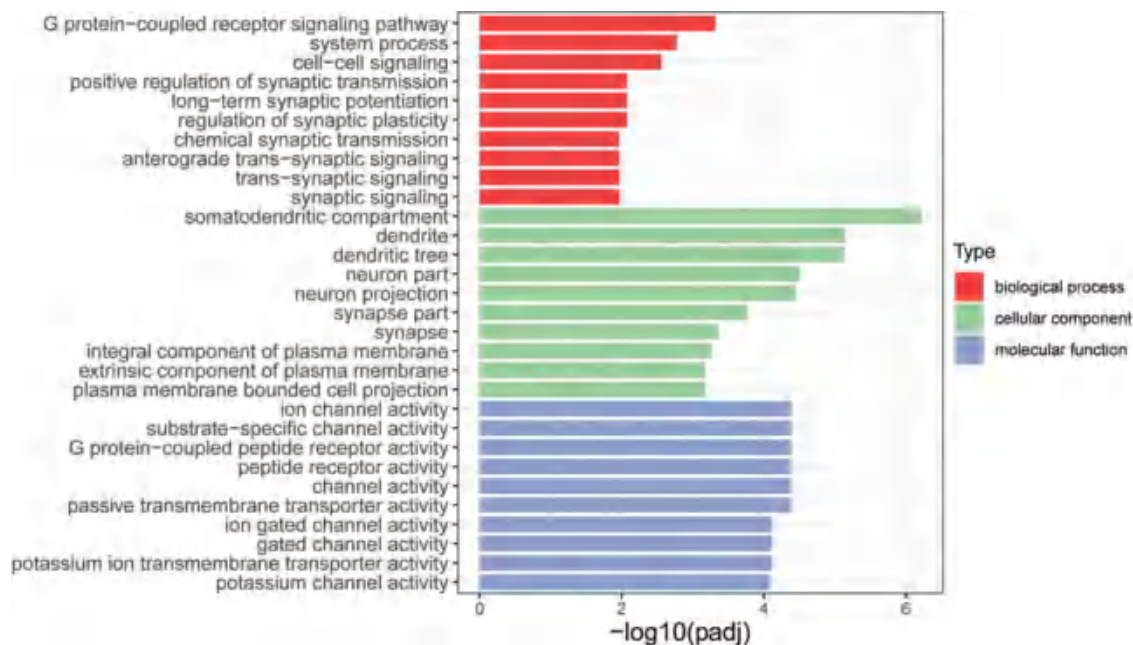
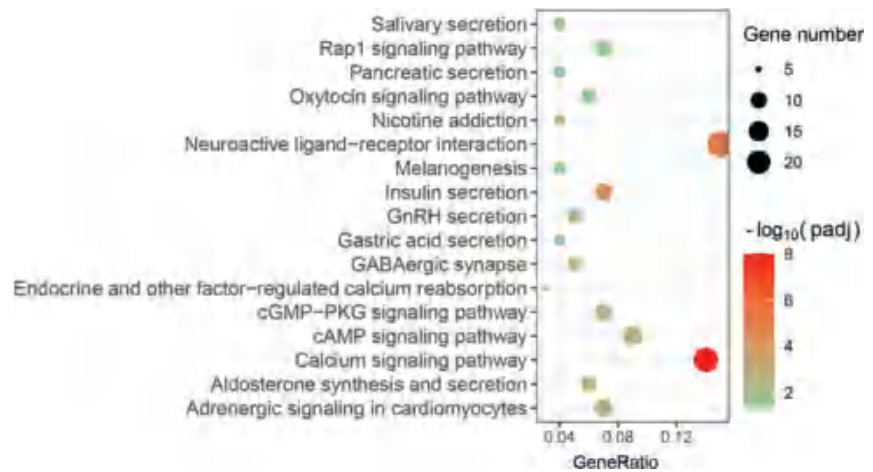


FIG. 3. GO enrichment analysis of the DEGs in the HE- and LE-groups. The DEGs were classified into 100 subcategories under the three GO categories: BP, MF, and CC. The barplot only displays the top 10 significantly enriched GO terms for three categories, and the remainder are shown in Supplementary Table S5. The x-axis represents the significance (adjusted p-value) of each GO term, and the y-axis represents each detailed classification of the GO term. BP, biological process; CC, cellular component; GO, Gene Ontology; MF, molecular function. Color images are available online.

FIG. 4. KEGG pathway enrichment analysis of the DEGs in the HE- and LE-groups. The DEGs were significantly enriched in 18 KEGG pathways. The x-axis indicates the proportion of DEGs enriched in a KEGG pathway to the total number of DEGs enriched in all KEGG pathways. The size of dot represents the number of DEGs enriched in each KEGG pathway, whereas the color of dot represents the significance (adjusted p-value) of each KEGG pathway. KEGG, Kyoto Encyclopedia of Genes and Genomes. Color images are available online.



TFs had positive values. Among 18 TFs, 2 of them showed extreme values for RIF1 or RIF2. POU6F2 showed the most extreme score for RIF1 (-2.49 SD), whereas TFAP2D had the most extreme score for RIF2 (-2.62 SD).

Gene co-expression network analysis

To indicate hub TFs related to FE of pigs, a co-expression correlation network ($|r| > 0.95$) between 18 differentially expressed TFs and their target DEGs was constructed according to the PCIT results (Supplementary Table S8). The co-expression network was composed of 11 TFs, 112 DEGs, and 212 connections (Fig. 5). The other seven TFs (RORB, RARB, SCRT1, ETV4, ESRRG, EMX1, and MEF2A) had no connection to DEGs ($|r| < 0.95$). All genes in the co-expression network were prioritized by their degree of connectivity, and TFs with a high degree were defined as hub TFs. The top five hub TFs were NR2F2, HES3, CEBPD, HNF1B, and DRGX, which vary from 21 to 50 degrees.

RT-qPCR Validation of DEGs Identified from RNA-seq

To validate the results of RNA-seq, a total of six DEGs (CABP1, COL11A2, EFNB3, GPR1, SCN1B, and SCN4B) were randomly selected for RT-qPCR. As shown in Figure 6, CABP1, COL11A2, GPR1, SCN1B, and SCN4B reached a significant level of difference (p -value < 0.05). Although the p -value of EFNB3 was more than 0.05 in the result of RT-qPCR, it had the same pattern of expression as observed with the RNA-seq. Therefore, gene expression observed in the brain transcriptome of eight pigs was credible.

Discussion

The improvement of the FE is a major priority in the porcine industry. Thus, there is a requirement to give deeper insights into essential genes and biological pathways that regulate the FE in pigs. In the current study, the phenotypic data showed that the HE-group had significantly lower feed intake compared with the LE-group, whereas there were no significant differences between HE- and LE-groups for ADG and AWMB, which indicated that the HE-group with low RFI have reduced feed costs compared with the LE-group without compromising growth. Therefore, based on

the crucial role of the brain in regulating food intake and energy homeostasis, brain transcriptomic analysis of high and low RFI pigs was performed to indicate the potential genes and biological pathways related to FE in pigs. As a result, we indicated a number of key TFs (TFAP2D, NR2F2, and HNF1B) and biological pathways (insulin secretion and Oxytocin signaling pathway) associated with FE of pigs.

The KEGG pathway enrichment analysis showed that the calcium signaling pathway and neuroactive ligand-receptor interaction are the most significant pathways. Previous studies have shown that these two pathways are essential for diverse cellular functions in the brain (Berridge, 2014; Dan et al., 2018). Moreover, these two pathways are also directly related to FE traits. Previous studies showed that the calcium signaling pathway was significantly associated with RFI in cattle (Rolf et al., 2012) and the neuroactive ligand-receptor interaction might play an important role in the regulation of feed intake in chickens (Cao et al., 2020; Li et al., 2018). Our results indicate that the effects of these two pathways on FE are applicable in a variety of animals, including pigs.

The GO and KEGG pathway enrichment analysis also showed that the DEGs were mainly enriched in synaptic signaling or transmission, and hormone secretion, in which insulin secretion and oxytocin signaling pathways were closely related to FE in pigs.

Insulin is an important metabolic hormone that promotes a range of growth processes through their central actions in the brain (Benoit et al., 2013; Fuller and Crofts, 1977). Insulin receptors are widely distributed in the brain, including the mesolimbic system of the ventral tegmental area and the arcuate nucleus of the hypothalamus, which recognize insulin to inhibit feed intake and regulate energy homeostasis (Davis et al., 2010a; Gerozissis, 2003). A previous study found that the allele frequency of RFI selection lines differed in genes related to insulin release (Oteru et al., 2013). Similarly, another previous study showed that DEGs related to insulin were over-expressed in pigs from the low RFI (Gondret et al., 2017). In our study, the majority of DEGs enriched in the insulin secretion pathway were upregulated in the HE-group, indicating that the secretion of insulin was more activated in the HE-group, which was consistent with previous studies and our result that the HE-group consumed less feed. Moreover, the insulin stimulates protein synthesis in the muscle tissues of pigs (Davis et al., 2010b; Davis

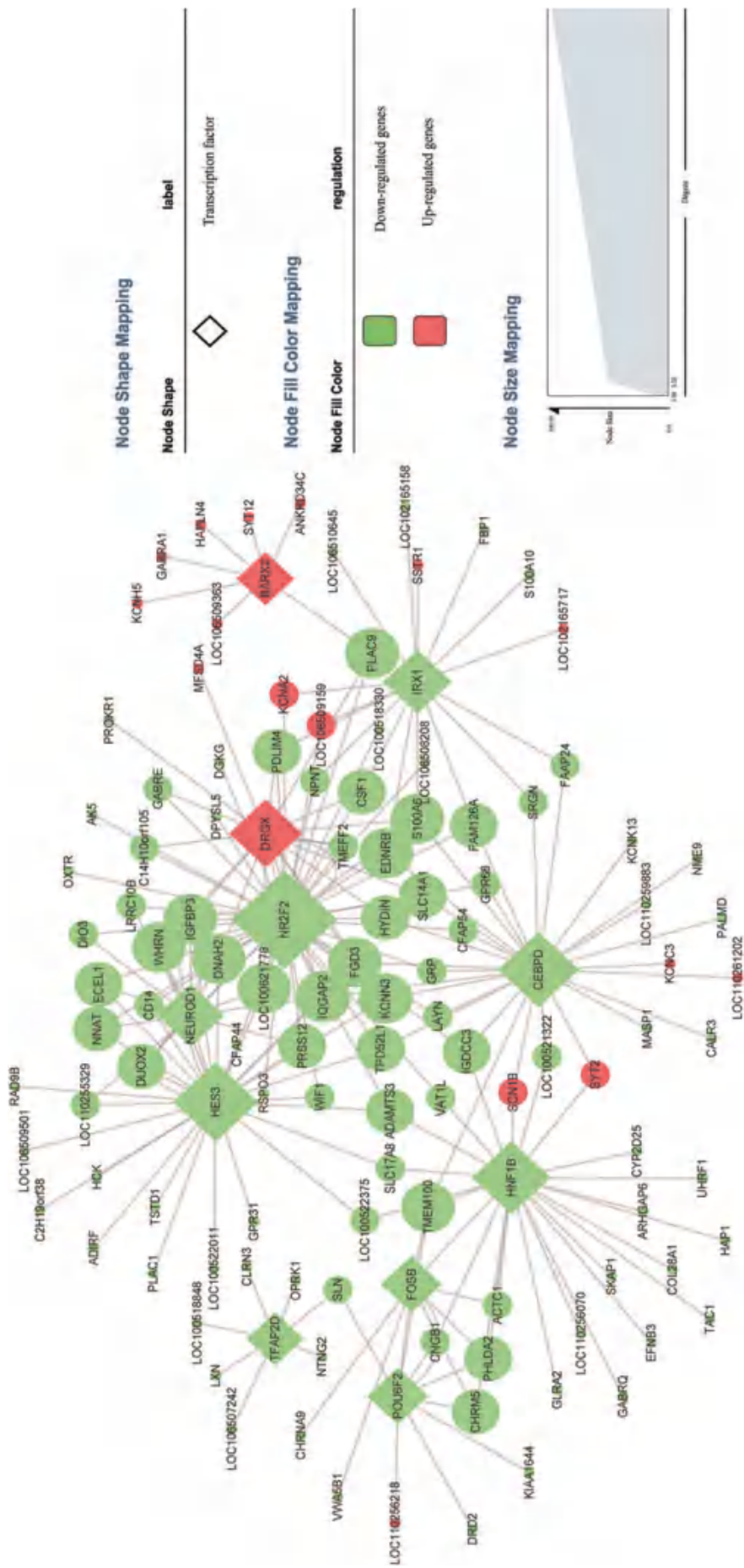


FIG. 5. Gene co-expression network of differentially expressed TFs and their respective DEGs with significant correlations above 0.95. The co-expression network was composed of 11 TFs, 112 DEGs, and 212 connections. Nodes with square shape were TFs, whereas those with circle shape were DEGs. Nodes filled with green were up-regulated genes, and those filled with red were down-regulated genes. The size of the nodes represents the degrees of the DEGs or TFs. Color images are available online.

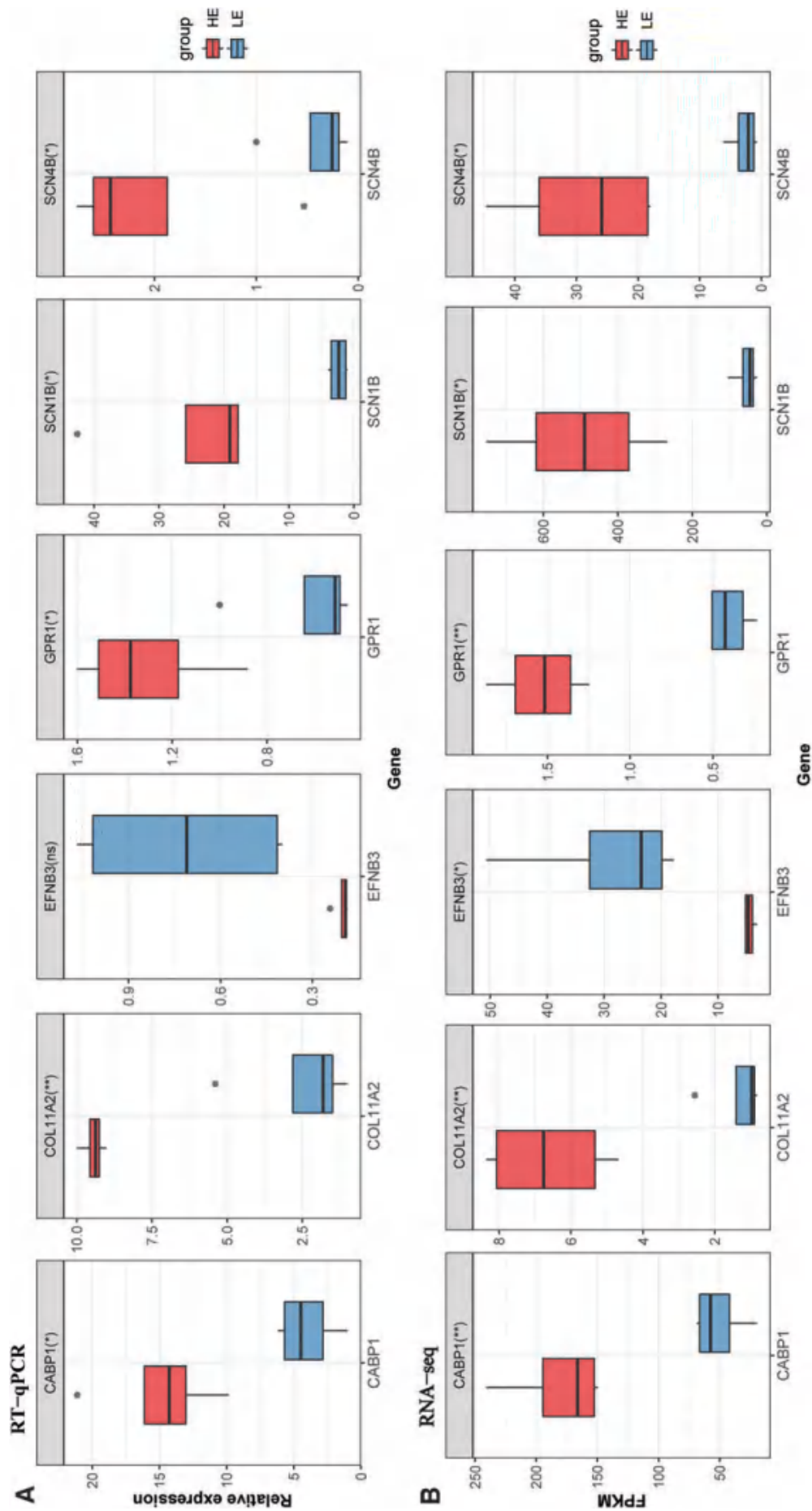


FIG. 6. Validation of six randomly selected DEGs by RT-qPCR. (A) The relative expression levels of six randomly selected DEGs (CABP1, COL11A2, EFNB3, GPR1, SCN1B, and SCN4B). The x-axis represents the gene name, whereas the y-axis represents the relative expression levels of these DEGs. *p-Value <0.05 and **p-value <0.01 in the HE-group compared with the LE-group. Red and blue color in the boxplot represent the HE-group and LE-group, respectively. (B) Boxplot of FPKM expression values for the six DEGs in RNA-seq. The x-axis represents the gene name, whereas the y-axis represents the FPKM expression levels of these DEGs. FPKM, fragment per kilobases of transcript per million reads; RNA-seq, real-time quantitative polymerase chain reaction. Color images are available online.

et al., 2002; O'Connor et al., 2003) and muscle growth seems to be associated with high FE animals, which might result in higher FE in the HE-group.

Our result also reveals that the oxytocin signaling pathway is associated with FE and most of the DEGs enriched in it are upregulated. Likewise, oxytocin is a well-characterized neurotransmitter that potentially affects FE, which has been suggested to be involved in a range of BPs, including energy metabolism and suppression of food intake (Alexandre et al., 2019; Takayanagi et al., 2008). Simultaneously, the oxytocin induces a slight change in leptin, which affects energy metabolism (Xu et al., 2011).

As a result, we propose that the reduction of feed intake in the HE-group might result from the insulin and oxytocin release. Moreover, it might be a difference in energy metabolism between two groups so that the HE-group has reduced feed intake but without compromising growth.

Further, the TFs analysis and gene co-expression network might provide insights into the potential drivers of transcriptional changes between high and low FE pigs in the brain.

The TFs analysis revealed that TFAP2D had the highest RIF scores ($RIF2 = -2.61SD$) among all TFs, suggesting that TFAP2D might be responsible for the differences in gene expression observed between the two groups. Transcription factor AP-2 delta (TFAP2D) is a protein-coding gene that is one of the members of the transcription factor activator protein-2 (TFAP2) family (Zhao et al., 2001). TFAP2 regulated genes involved in the process of development, cell growth, differentiation, and apoptosis (Kristina et al., 2000). TFAP2 can also initiate feeding behavior and is higher expressed in the LE-group, which is consistent with our results that the LE-group consumed more feed. A previous study showed that the flies with overexpression of TFAP2 consumed more food and had threefold as many feeding bouts as controls (Goergen, 2014). Moreover, high-RFI and low-RFI pigs showed significant differences in allele frequencies of TFAP2A (one of the members of TFAP2 family) in a whole-genome association study (Onteru et al., 2013). Combining with the previous study, these results suggest that TFAP2D might be a hub gene regulated FE in pigs.

The gene co-expression network analysis combining information obtained from PCIT and RIF analysis showed that NR2F2 appeared as the most connected TFs and was downregulated in the HE-group. Nuclear receptor subfamily 2 group F member 2 (NR2F2) is a member of the steroid/thyroid hormone receptor superfamily and is also named chicken ovalbumin upstream promoter transcription factor II (COUP-TFII), which plays a pivotal role in glucose homeostasis and energy metabolism (Nobuyuki et al., 1988; Li et al., 2009). NR2F2 is associated with whole-body insulin sensitivity and is directly repressed by the exogenous insulin (Perilhou et al., 2008; Marie et al., 2012), which is consistent with our study that the LE-group secretes less insulin than the HE-group. The NR2F-family also includes two orphan receptors: NR2F1 and NR2F6 (Hermann-Kleiter and Baier, 2014). A previous study showed that NR2F6 was a key regulator associated with FE in beef cattle (Alexandre et al., 2019). Therefore, previous and current studies reveal an important role of NR2F2 in regulating FE.

Another TF prioritized in our study was HNF1B (hepatocyte nuclear factor 1B), which was the fourth most connected gene in the co-expression network and was downregulated in

the HE-group. Consistent with our result, HNF1B has previously been identified as DEGs in cattle divergent for RFI and were found nearby to a variant associated with feed intake (Xi et al., 2015), and this gene is also identified as an upregulated gene in low-FE cattle (Al-Husseini et al., 2016). As a result, NR2F2, and HNF1B might be considered crucial genes that affect FE in pigs.

Conclusion

This study was the first brain transcriptomic analysis conducted on pigs that diverged in RFI to identify the TFs and biological pathways related to FE. By comparing phenotypic data between HE- and LE-groups, we found that the HE-group had significantly lower feed intake but without compromising growth. A total of 430 DEGs were identified between two groups, among which 18 TFs were found. From the results of KEGG analysis, insulin secretion and the oxytocin signaling pathway were closely related to FE by reducing feed intake and regulating energy metabolism. Importantly, three hub TFs (NR2F2, TFAP2D, and HNF1B) were identified in TFs analysis and gene co-expression analysis, which might regulate FE mainly by affecting feeding behavior, insulin sensitivity, or regulating energy metabolism.

Briefly, our results indicated that the HE-group may have increased insulin and oxytocin secretion to reduce food intake and maintain normal energy metabolism without compromising growth, thereby showing higher FE. Our results provide new insights into the molecular mechanisms in the brain in regulating FE and these findings potentially help in improving FE in pigs.

Disclosure Statement

The authors have declared that no competing interests exist.

Funding Information

This study was supported by the Local Innovative and Research Teams Project of Guangdong Province (2019BT02 N630), the Natural Science Foundation of Guangdong Province (2018B030315007), and the Pearl River S and T Nova Program of Guangzhou (201906010011).

Supplementary Material

Supplementary Table S1
Supplementary Table S2
Supplementary Table S3
Supplementary Table S4
Supplementary Table S5
Supplementary Table S6
Supplementary Table S7
Supplementary Table S8

References

Al-Husseini, W., Chen, Y., Gondro, C., Herd, R.M., Gibson, J.P., and Arthur, P.F. (2016). Characterization and profiling of liver microRNAs by RNA-sequencing in cattle divergently selected for residual feed intake. *Asian-Australas J Anim Sci* 29, 1371–1382.

- Alexander, D., Davis, C.A., Felix, S., Jorg, D., Chris, Z., Sonali, J., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Alexandre, P.A., Naval-Sanchez, M., Porto-Neto, L.R., Ferraz, J.B.S., Reverter, A., and Fukumasu, H. (2019). Systems biology reveals NR2F6 and TGFB1 as key regulators of feed efficiency in beef cattle. *Front Genet* 10, 230.
- Baskin, D.G., Brewitt, B., Davidson, D.A., Corp, E., Paquette, T., Figlewicz, D.P., et al. (1986). Quantitative autoradiographic evidence for insulin receptors in the choroid plexus of the rat brain. *Diabetes* 35, 246–249.
- Benoit, S.C., Clegg, D.J., Seeley, R.J., and Woods, S.C. (2013). Insulin and leptin as adiposity signals. *Recent Prog Horm Res* 59, 267–285.
- Berridge MJ. (2014). Calcium signalling and psychiatric disease: bipolar disorder and schizophrenia. *Cell Tissue Res* 357, 477–492.
- Berthoud, H.R. (2002). Multiple neural systems controlling food intake and body weight. *Neurosci Biobehav R* 26, 393–428.
- Blevins, J.E., Schwartz, M.W., and Baskin, D.G. (2002). Peptide signals regulating food intake and energy homeostasis. *Can J Physiol Pharmacol* 80, 396–406.
- Boddicker, N., Gabler, N., Spurlock, M., Nettleton, D., and Dekkers, J. (2011). Effects of ad libitum and restricted feed intake on growth performance and body composition of Yorkshire pigs selected for reduced residual feed intake. *J Anim Sci* 89, 40–51.
- Cai, W., Casey, D., and Dekkers, J. (2008). Selection response and genetic parameters for residual feed intake in Yorkshire swine. *J Anim Sci* 86, 287–298.
- Cao, X., Wang, Y., Shu, D., Qu, H., Luo, C., and Hu, X. (2020). Food intake-related genes in chicken determined through combinatorial genome-wide association study and transcriptome analysis. *Anim Genet* 51, 741–751.
- Dan, M., Wen, H., Shao, A., and Xu, L. (2018). Silver nanoparticle exposure induces neurotoxicity in the rat hippocampus without increasing the blood-brain barrier permeability. *J Biomed Nanotechnol* 14, 1330–1338.
- Davis, J.F., Choi, D.L., and Benoit, S.C. (2010a). Insulin, leptin and reward. *Trends Endocrinol Metab* 21, 68–74.
- Davis, T.A., Fiorotto, M.L., Burrin, D.G., Reeds, P.J., Nguyen, H.V., Beckett, P.R., et al. (2002). Stimulation of protein synthesis by both insulin and amino acids is unique to skeletal muscle in neonatal pigs. *Am J Physiol Endocrinol Metab* 282, E880–E890.
- Davis, T.A., Suryawan, A., Orellana, R.A., Fiorotto, M.L., and Burrin, D.G. (2010b). Amino acids and insulin are regulators of muscle protein synthesis in neonatal pigs. *Animal* 4, 1790–1796.
- De Biagi, C.A.O., Nociti, R.P., Funiceli, B.O., de Cássia Ruy, P., Ximenez, J.P.B., and Silva, W.A. (2020). CeTF: an R package to coexpression for transcription factors using regulatory impact factors (RIF) and partial correlation and information (PCIT) analysis. *bioRxiv*; DOI: 10.1101/2020.03.30.015784.
- Fuller, M.F., and Crofts, R.M. (1977). The protein-sparing effect of carbohydrate. *Br J Nutr* 38, 489–496.
- Gerozissis, K. (2003). Brain insulin: regulation, mechanisms of action and functions. *Cell Mol Neurobiol* 23, 1–25.
- Gilbert, H., Bidanel, J.-P., Gruand, J., Caritez, J.-C., Billon, Y., Guillouet, P., et al. (2007). Genetic parameters for residual feed intake in growing pigs, with emphasis on genetic relationships with carcass and meat quality traits. *J Anim Sci* 85, 3182–3188.
- Goergen, P. (2014). The Molecular Mechanism of Aggression and Feeding Behaviour in *Drosophila melanogaster*. Uppsala University.
- Gondret, F., Vincent, A., Houée-Bigot, M., Siegel, A., Lagarrigue, S., Causeur, D., et al. (2017). A transcriptome multi-tissue analysis identifies biological pathways and genes associated with variations in feed efficiency of growing pigs. *BMC Genomics* 18, 244.
- Hu, H., Miao, Y., Jia, L., Yu, Q., Zhang, Q., and Guo, A. (2019). AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res* 47, D33.
- Hermann-Kleiter, N., and Baier, G. (2014). Orphan nuclear receptor NR2F6 acts as an essential gatekeeper of Th17 CD4+T cell effector functions. *Cell Commun Signal* 12, 38.
- Hilger-Eversheim, K., Moser, M., Schorle, H., and Buettner, R. (2000). Regulatory roles of AP-2 transcription factors in vertebrate development, apoptosis and cell-cycle control. *Gene* 260, 1–12.
- Hoque, M.A., and Suzuki, K. (2009). Genetics of residual feed intake in cattle and pigs: a review. *Asian-Australas J Anim Sci* 22, 747–755.
- Horodyska, J., Wimmers, K., Reyer, H., Trakooljul, N., Mullen, A.M., Lawlor, P.G., et al. (2018). RNA-seq of muscle from pigs divergent in feed efficiency and product quality identifies differences in immune response, growth, and macronutrient and connective tissue metabolism. *BMC Genomics* 19, 791.
- Jing, L., Hou, Y., Wu, H., Miao, Y., Li, X., Cao, J., et al. (2015). Transcriptome analysis of mRNA and miRNA in skeletal muscle indicates an important network for differential Residual Feed Intake in pigs. *Sci Rep* 5, 11953.
- Koch, R.M., Swiger, L.A., Chambers, D., and Gregory, K.E. (1963). Efficiency of feed use in beef cattle. *J Anim Sci* 22, 486–494.
- Li, L., Xin, X., Qin, J., Jeha, G.S., Saha, P.K., Yan, J., et al. (2009). The nuclear orphan receptor COUP-TFII plays an essential role in adipogenesis, glucose homeostasis, and energy metabolism. *Cell Metab* 9, 77–87.
- Li, Z., Liu, X., Zhang, P., Han, R., Sun, G., Jiang, R., et al. (2018). Comparative transcriptome analysis of hypothalamus-regulated feed intake induced by exogenous visfatin in chicks. *BMC Genomics* 19, 249.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550.
- Marie, B., Pereira, R.O.H., Cécile, L., Emmanuel, V., Julien, P., Zhang, P., et al. (2012). Glucose-dependent regulation of NR2F2 promoter and influence of SNP-rs3743462 on whole body insulin sensitivity. *PLoS One* 7, p1–p9.9p.
- Mcminn, J.E., Baskin, D.G., and Schwartz, M.W. (2000). Neuroendocrine mechanisms regulating food intake and body weight. *Obes Rev* 1, 37–46.
- Nobuyuki, M., Yasunori, K., Shin-ichi, F., Sho-ichi, S., Kentaro, S., Yuji, Y., et al. (1988). Identification of two novel members of erbA superfamily by molecular cloning: the gene products of the two are highly related to each other. *Nucleic Acids Res* 16, 11057–11074.
- Obici, S., and Rossetti, L. (2003). Minireview: nutrient sensing and the regulation of insulin action and energy balance. *Endocrinology*, 144, 5172–5178.
- Obici, S., Zhang, B.B., Karkanas, G., and Rossetti, L. (2002). Hypothalamic insulin signaling is required for inhibition of glucose production. *Nat Med* 8, 1376–1382.
- O'Connor, P.M., Bush, J.A., Suryawan, A., Nguyen, H.V., and Davis, T.A. (2003). Insulin and amino acids independently stimulate skeletal muscle protein synthesis in neonatal pigs. *Am J Physiol Endocrinol Metab* 284, E110–E119.

- Oczkowicz, M., Szmatoła, T., Piórkowska, K., and Ropka-Molik, K. (2018). Variant calling from RNA-seq data of the brain transcriptome of pigs and its application for allele-specific expression and imprinting analysis. *Gene* 641, 367–375.
- Onteru, S.K., Gorbach, D.M., Young, J.M., Garrick, D.J., Dekkers, J.C.M., Rothschild, M.F., and Liu, Z. (2013). Whole genome association studies of residual feed intake and related traits in the pig. *PLoS One* 8, e61756.
- Patience, J.F., Rossoni-Serão, M.C., and Gutiérrez, N.A. (2015). A review of feed efficiency in swine: biology and application. *J Anim Sci Biotechnol* 6, 33.
- Perillou, A., Tourrel-Cuzin, C., Kharroubi, I., Henique, C., Fauveau, V., Kitamura, T., et al. (2008). The transcription factor COUP-TFII is negatively regulated by insulin and glucose via Foxo1- and ChREBP-controlled pathways. *Mol Cell Biol* 28, 6568–6579.
- Rakhshandeh, A., Dekkers, J.C., Kerr, B.J., Weber, T.E., English, J., and Gabler, N.K. (2012). Effect of immune system stimulation and divergent selection for residual feed intake on digestive capacity of the small intestine in growing pigs. *J Anim Sci* 90 Suppl 4, 233–235.
- Reverter, A., Hudson, N.J., Nagaraj, S.H., Pérez-Enciso, M., and Dalrymple, B.P. (2010). Regulatory impact factors: unraveling the transcriptional regulation of complex traits from expression data. *Bioinformatics* 26, 896–904.
- Rolf, M.M., Taylor, J.F., Schnabel, R.D., McKay, S.D., McClure, M.C., Northcutt, S.L., et al. (2012). Genome-wide association analysis for feed efficiency in Angus cattle. *Anim Genet* Aug 43, 367–374.
- Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 3, 1101–1108.
- Shannon, P., Markiel, A., Ozier, O., Baliga, S.N., Wang, T.J., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13, 2498–2504.
- Simon, A., Theodor, P.P., and Wolfgang, H. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169.
- Takayanagi, Y., Kasahara, Y., Onaka, T., Takahashi, N., Kawada, T., and Nishimori, K. (2008). Oxytocin receptor-deficient mice developed late-onset obesity. *Neuroreport* 19, 951–955.
- Teagasc. (2018). Pig Herd Performance Report 2018. Foster City, CA: Teagasc Pig Development Department.
- Trayhurn, P., Bing, C., and Wood, I.S. (2006). Adipose tissue and adipokines—energy regulation from the human perspective. *J Nutr* 136, 1935S–1939S.
- Wang, X., Li, S., Wu, J., Ding, R., Quan, J., Zheng, E., et al. (2019). A transcriptome analysis identifies biological pathways and candidate genes for feed efficiency in DLY pigs. *Genes* 10, 725.
- Woods, S.C., Seeley, R.J., Porte, D., Jr., and Schwartz, M.W. (1998). Signals that regulate food intake and energy homeostasis. *Science* 280, 1378–1383.
- Xi, Y.M., Yang, Z., Wu, F., Han, Z.Y., and Wang, G.L. (2015). gene expression profiling of hormonal regulation related to the residual feed intake of holstein cattle. *Biochem Biophys Res Commun* 465, 19–25.
- Xu, C., Wang, X., Zhuang, Z., Wu, J., Zhou, S., Quan, J., et al. (2020). A transcriptome analysis reveals that hepatic glycolysis and lipid synthesis are negatively associated with feed efficiency in DLY pigs. *Sci Rep* 10, 1–12.
- Xu, Y., Elmquist, J.K., and Fukuda, M. (2011). Central nervous control of energy and glucose balance: focus on the central melanocortin system. *Ann N Y Acad Sci* 1243, 1–14.
- Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). ClusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287.
- Zhao, F., Satoda, M., Licht, J.D., Hayashizaki, Y., and Gelb, B.D. (2001). Cloning and characterization of a novel mouse AP-2 transcription factor, AP-2delta, with unique DNA binding and transactivation properties. *J Biol Chem* 276, 40755.

Address correspondence to:

Enqin Zheng

College of Animal Science and National Engineering
Research Center for Breeding Swine Industry
South China Agricultural University
Guangdong 510642
P.R. China

E-mail: eqzheng@scau.edu.cn

Jie Yang, PhD

College of Animal Science and National Engineering
Research Center for Breeding Swine Industry
South China Agricultural University
Guangdong 510642
P.R. China

E-mail: jieyang2012@hotmail.com

Received for publication August 23, 2020; received in revised form October 27, 2020; accepted November 3, 2020.



广东省科技进步奖 证书

为表彰 2023 年度广东省科技
进步奖获得者，特颁发此证书。

项目名称: 瘦肉型种猪基因组育种技术与应用

奖励等级: 一等奖

获奖者: 郑恩琴

粤府证【2024】1795 号
项目编号: J021-1-01-R10



高等学校科学研究优秀成果奖 (科学技术)

证 书

项目名称: 新型高效“温氏 WS501 猪配套系”培育
与应用

奖励类别: 科学技术进步奖

奖励等级: 一等奖

获 奖 者: 郑恩琴



证书编号: 2022-407-R010

神农中华农业科技奖 证书

为表彰在我国农业科学技术进步工作中做出突出贡献的获奖者，特颁发此证书，以资鼓励。

成果名称：五系配套瘦肉型猪选育关键技术与应用
奖励等级：一等奖
获奖者：郑恩琴（第13完成人）
获奖者单位：华南农业大学

证书编号：2021-KJ035-1-R13





广东省农业技术推广奖

获奖证书

为表彰在农业技术推广工作中做出贡献的单位和个人，特颁发此证书，以资鼓励。

获奖项目：五元杂交瘦肉型种猪新配套系及其养殖技术示范推广

奖励等级：一等奖

获奖者：郑恩琴

奖励日期：2020年12月17日

证书号：2019-1-X02-R12



证书号第 15797722 号



实用新型专利证书

实用新型名称：一种用于肉质样品测定预处理的恒温水浴装置

发 明 人：郑恩琴;李颢;霍梦飞;张学政;张航瑞

专 利 号：ZL 2021 2 1736846.3

专利申请日：2021 年 07 月 28 日

专 利 权 人：华南农业大学

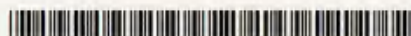
地 址：510640 广东省广州市天河区五山路 483 号

授权公告日：2022 年 02 月 11 日

授权公告号：CN 215783481 U

国家知识产权局依照中华人民共和国专利法经过初步审查，决定授予专利权，颁发实用新型专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为十年，自申请日起算。

专利证书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第 1 页 (共 2 页)

其他事项参见续页

证书号第6136934号



发明专利证书

发明名称：一种猪1号染色体上与猪剩余采食量相关的SNP分子标记及其用途

发明人：杨杰;吴珍芳;丁荣荣;郑恩琴;吴杰;蔡更元;洪林君
杨化强;黄思秀

专利号：ZL 2021 1 0842930.1

专利申请日：2021年07月26日

专利权人：华南农业大学

地址：510642 广东省广州市天河区五山路483号

授权公告日：2023年07月14日

授权公告号：CN 113699247 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第1页(共2页)

其他事项参见续页

证书号第6130519号



发明专利证书

发明名称：一种影响猪饲料转化效率性状的SNP分子标记及其用途

发明人：吴珍芳;杨杰;丁荣荣;郑恩琴;蔡更元;李紫聪;吴杰
杨化强;洪林君;黄思秀

专利号：ZL 2021 1 0842685.4

专利申请日：2021年07月26日

专利权人：华南农业大学

地址：510642 广东省广州市天河区五山路483号

授权公告日：2023年07月11日

授权公告号：CN 113699246 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



证书号第 5368757 号



发明专利证书

发明名称：一种猪 11 号染色体上影响猪背膘厚的拷贝数变异分子标记及应用

发明人：杨杰;吴珍芳;阮栋林;郑恩琴;丁荣荣;邱益彬;庄站伟
吴杰;顾婷;洪林君;徐铮;黄思秀

专利号：ZL 2020 1 1375352.7

专利申请日：2020 年 11 月 30 日

专利权人：华南农业大学

地址：510640 广东省广州市天河区五山路 483 号

授权公告日：2022 年 08 月 09 日

授权公告号：CN 112280874 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利证书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第 1 页 (共 2 页)

其他事项参见续页

证书号第5101252号



发明专利证书

发明名称：一种与猪剩余采食量相关的 SNP 分子标记及其用途

发明人：吴珍芳;丁荣荣;杨杰;蔡更元;郑恩琴;庄站伟;李紫聪
徐铮;顾婷

专利号：ZL 2021 1 0842691.X

专利申请日：2021 年 07 月 26 日

专利权人：温氏食品集团股份有限公司;华南农业大学

地址：527400 广东省云浮市新兴县新城镇东堤北路 9 号

授权公告日：2022 年 04 月 22 日

授权公告号：CN 113584181 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利证书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第 1 页 (共 2 页)

其他事项参见续页

证书号第 5496896 号



发明专利证书

发明名称：一种位于猪 14 号染色体上与母猪死胎数和健仔率相关的 SNP 分子标记及其用途

发明人：吴珍芳;杨杰;周身娉;丁荣荣;郑恩琴;蔡更元;全建平
李紫聪;顾婷;徐铮

专利号：ZL 2021 1 0962528.7

专利申请日：2021 年 08 月 20 日

专利权人：华南农业大学

地址：510642 广东省广州市天河区五山路 483 号

授权公告日：2022 年 10 月 04 日

授权公告号：CN 113564264 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利证书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第 1 页 (共 2 页)

其他事项参见续页

证书号第6135778号



发明专利证书

发明名称：一种猪13号染色体上与母猪产畸形仔猪数相关的SNP分子标记及其用途

发明人：杨杰;吴珍芳;邱益彬;丁荣荣;郑恩琴;蔡更元;李紫聪
洪林君;徐铎

专利号：ZL 2021 1 0905687.3

专利申请日：2021年08月06日

专利权人：华南农业大学

地址：510642 广东省广州市天河区五山路483号

授权公告日：2023年07月14日

授权公告号：CN 113637768 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第1页(共2页)

其他事项参见续页

证书号第6128885号



发明专利证书

发明名称：一种猪7号染色体上与猪死胎数、活仔率相关的SNP分子标记及其用途

发明人：杨杰;吴珍芳;阮栋林;丁荣荣;郑恩琴;全建平;蔡更元
洪林君;黄思秀;杨化强

专利号：ZL 2021 1 0873960.9

专利申请日：2021年07月30日

专利权人：华南农业大学

地址：510642 广东省广州市天河区五山路483号

授权公告日：2023年07月11日

授权公告号：CN 113736889 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第1页(共2页)

其他事项参见续页

证书号第5477463号



发明专利证书

发明名称：一种猪3号染色体上与猪日增重和上市体重日龄相关的拷贝数变异分子标记及应用

发明人：蔡更元;吴珍芳;杨杰;陈悦;郑恩琴;庄站伟;丁荣荣
邱益彬;吴杰;杨化强;洪林君;黄思秀

专利号：ZL 2020 1 1375366.9

专利申请日：2020年11月30日

专利权人：华南农业大学

地址：510640 广东省广州市天河区五山路483号

授权公告日：2022年09月27日

授权公告号：CN 112458183 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利证书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第1页(共2页)

其他事项参见续页

证书号第 4036343 号



发明专利证书

发明名称：位于猪 16 号染色体上与猪瘦肉率和眼肌面积相关的 SNP 分子标记及应用

发明人：杨杰;吴珍芳;周身娉;蔡更元;郑恩琴;徐铮

专利号：ZL 2019 1 0403266.3

专利申请日：2019 年 05 月 15 日

专利权人：华南农业大学

地址：510642 广东省广州市天河区五山路 483 号

授权公告日：2020 年 10 月 16 日

授权公告号：CN 110117665 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利证书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第 1 页 (共 2 页)

其他事项参见续页

证书号第 4024110 号



发明专利证书

发明名称：位于猪 7 号染色体上与总乳头数相关的 SNP 分子标记及应用

发明人：吴珍芳;杨杰;叶勇;庄站伟;郑恩琴;徐铮;蔡更元

专利号：ZL 2019 1 0407561.6

专利申请日：2019 年 05 月 15 日

专利权人：华南农业大学

地址：510642 广东省广州市天河区五山路 483 号

授权公告日：2020 年 10 月 09 日

授权公告号：CN 110144408 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利证书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第 1 页 (共 2 页)

其他事项参见续页

证书号第5283738号



发明专利证书

发明名称：猪6号染色体上与瘦肉率、眼肌面积、眼肌厚度相关的SNP分子标记与应用

发明人：吴珍芳;杨杰;庄站伟;蔡更元;郑恩琴;徐铮

专利号：ZL 2019 1 0385368.7

专利申请日：2019年05月09日

专利权人：华南农业大学

地址：510642 广东省广州市天河区五山路483号

授权公告日：2022年07月05日

授权公告号：CN 110257529 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利证书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第1页(共2页)

其他事项参见续页

证书号第 4666883 号



发明专利证书

发明名称：位于猪 7 号染色体上与杜洛克猪日增重性状相关的分子标记及应用

发明人：吴珍芳；杨杰；吴杰；周身娉；郑恩琴；徐铮；蔡更元

专利号：ZL 2019 1 0251984.3

专利申请日：2019 年 03 月 29 日

专利权人：华南农业大学

地址：510642 广东省广州市天河区五山路 483 号

授权公告日：2021 年 09 月 07 日

授权公告号：CN 110106255 B

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，颁发发明专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为二十年，自申请日起算。

专利证书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。



局长
申长雨

申长雨



第 1 页 (共 2 页)

其他事项参见续页

荣誉证书

郑恩琴：

在华南农业大学动物科学学院 2020 年青年教师
教学观摩比赛中表现突出，荣获
一等奖

特发此证，以资鼓励。

证书编号：DKQJS-2020-1-03

华南农业大学动物科学学院

2021 年 1 月 25 日