网络出版时间:2026-05-11 DOI:10.13207/j.jnwafu.2026.11.002

# 鸡*GTF2A1*基因SNP位点筛选及其与产蛋性状的关联分析

李春苗,黄华云,王钱保,吴兆林,黄正洋,孔令琳,胡阳阳,赵振华
(江苏省家禽科学研究所,江苏 扬州 225125)

[摘 要] 【目的】解析通用转录因子ⅡA亚基1(general transcription factor ⅡA subunit 1,GTF2A1)基因对鸡产蛋性状的遗传效应,筛选与其显著相关的单核苷酸多态性(single nucleotide polymorphism,SNP)标记,为苏禽4号母系(M系)产蛋性能的分子标记辅助育种提供依据。【方法】以384只苏禽4号母系(M系)为试验对象,采用Sanger测序与Sequenom MassARRAY技术对*GTF2A1*基因进行SNP筛查。利用Haploview v4.1分析连锁不平衡(linkage disequilibrium,LD),基于一般线性模型(general linear model,GLM)进行SNP位点及单倍型与产蛋性状的关联分析,并通过MEME Suite预测SNP对转录因子结合的影响。用GOMO对*GTF2A1*基因SNP位点保守基序进行功能预测,【结果】共检出2个SNP位点:g.40334089G>C(位于5′ UTR,为颠换突变)和g.40341114A>G(位于外显子3,为同义突变)。2个位点均符合哈代-温伯格平衡($P>0.05$),且呈中度多态性。关联分析表明,g.40334089G>C位点GG基因型个体与g.40341114A>G位点AA基因型个体的开产日龄和300日龄产蛋数均显著优于其他基因型个体($P<0.05$)。2个SNP位点处于强连锁不平衡状态,形成H1(GA)、H2(CG)、H3(GG)等3种单倍型,其中H1H1(GGAA)组合个体的300日龄产蛋数与开产日龄表现较优。g.40341114A>G位点有一个保守基序(CCTGCAGCTGGTCAG-CAGCAGCCR(A/G)CAGC),该保守基序参与了小GTP酶介导的信号转导、蛋白质转运、葡萄糖代谢过程等生物过程及趋化因子活性、GTP结合等分子功能;g.40334089G>C位点没有保守基序。g.40341114A>G位点A→G突变能够增强保守基序与转录因子Myod1的结合能力,G→A突变可增强保守基序与转录因子REST和PPARG的结合能力。【结论】*GTF2A1*基因中g.40334089G>C和g.40341114A>G位点均与产蛋性状显著相关,其中单倍型组合H1H1(GGAA)为改善开产日龄和300日龄产蛋数的优势遗传标记,可作为苏禽4号母系(M系)分子标记辅助育种的核心依据。

[关键词] 鸡;*GTF2A1*基因;产蛋性状;单核苷酸多态性
[中图分类号] S831.2 [文献标志码] A [文章编号] 1671-9387(2026)11-0011-10

# Screening and bioinformatics analysis of SNPs in the chicken *GTF2A1* gene and their association with egg-laying traits

LI Chunmiao, HUANG Huayun, WANG Qianbao, WU Zhaolin, HUANG Zhengyang, KONG Linglin, HU Yangyang, ZHAO Zhenhua
(*Jiangsu Institute of Poultry Sciences*, *Yangzhou*, *Jiangsu* 225125, *China*)

**Abstract**: 【Objective】 This study aimed to elucidate the genetic effects of the general transcription factor ⅡA Subunit 1 (*GTF2A1*) gene on egg-laying traits in chickens, and screen single nucleotide polymorphism (SNP) markers significantly associated with these traits, to provide a basis for marker-assisted selection to improve egg-laying performance in the maternal line (M line) of Suqin No. 4 chickens. 【Method】 A total of 384

[收稿日期] 2025-09-12 [修回日期] 2025-10-16
[基金项目] 现代农业产业技术体系建设专项(CARS-41-Z21);江苏省种业振兴揭榜挂帅项目(JBGS(2021)109)
[作者简介] 李春苗(1982—),女,安徽砀山人,研究员,硕士,主要从事肉鸡育种与生产研究。E-mail:cherrylcm@163.com
[通信作者] 赵振华(1972—),男,江苏太仓人,研究员,博士,硕士生导师,主要从事肉鸡育种与生产研究。E-mail:zzh0514@163.com

hens from the Suqin No. 4 maternal line were used as experimental subjects. SNP screening of the *GTF2A1* gene was performed by using Sanger sequencing and Sequenom MassARRAY technology. Haploview v4.1 was employed for linkage disequilibrium (LD) analysis. Association analyses between SNP loci/haplotypes and egg-laying traits were conducted based on a general linear model (GLM). The potential impact of SNP on transcription factor binding was predicted by using the MEME Suite. Function prediction of conserved motifs for SNP loci in the *GTF2A1* gene was performed by using GOMO. 【Result】 Two SNP loci were identified: g.40334089G>C (located in the 5′ UTR, a transversion mutation) and g.40341114A>G (located in exon 3, a synonymous mutation). Both loci were in Hardy-Weinberg equilibrium ($P>0.05$) and exhibited moderate polymorphism. Association analysis revealed that individuals with the GG genotype at the g.40334089G>C locus and individuals with the AA genotype at the g.40341114A>G locus had significantly better performance in age at first egg and the total number of eggs laid at the age of 300 days, compared to individuals with other genotypes ($P<0.05$). The two SNP loci were in strong linkage disequilibrium, forming three haplotypes: H1(GA), H2(CG), and H3(GG), with H1H1 (GGAA) diplotype individuals showing better performance in age at first egg and the total number of eggs laid at the age of 300 days. The g.40341114A>G locus contained one conserved motif (CCTGCAGCTGGTCAGCAGCAGCCR(A/G)CAGC), which was involved in biological processes such as small GTPase-mediated signal transduction, protein transport, and glucose metabolic process, as well as molecular functions including chemokine activity and GTP binding. No conserved motif was identified in the g.40334089G>C locus. The A→G mutation at the g.40341114A>G locus could enhance the binding affinity of the conserved motif to the transcription factor Myod1, while the G→A mutation could enhance its binding affinity to transcription factors REST and PPARG. 【Conclusion】 The g.40334089G>C and g.40341114A>G loci in the *GTF2A1* gene are significantly associated with egg-laying traits. The H1H1 (GGAA) diplotype serves as a favorable genetic marker for improving age at first egg and the total number of eggs laid at the age of 300 days, and can be used as a core basis for marker-assisted selection in the maternal line (M line) of Suqin No. 4 chickens.

**Key words**: chicken; *GTF2A1* gene; egg-laying traits; single nucleotide polymorphism

种鸡产蛋性状是家禽遗传育种的核心指标,其遗传优化直接决定了种鸡的繁殖性能与产业的可持续发展能力。应用现代基因组学技术系统解析种鸡遗传特征,筛选关键分子标记,对提升种鸡繁殖性能具有重大的实践价值。美国国家生物技术信息中心官网公示的鸡通用转录因子ⅡA亚基1(general transcription factor ⅡA Subunit 1, GTF2A1)基因分布在5号染色体[1]上,长度为26.3 kb,由8个内含子和9个外显子组成。真核RNA聚合酶无法单独识别并结合启动子,而是依赖基本转录因子与启动子形成前起始复合物以启动转录,其中*GTF2A1*作为关键因子,对大多数启动子的转录激活是必需的[2]。GTF2A1蛋白可与TATA区域结合蛋白相互作用,在转录激活中发挥关键作用[3]。*GTF2A1*基因的功能不仅限于转录调控,其在生殖细胞发育中也发挥着重要作用,近年来,*GTF2A1*基因在动物生殖系统中的潜在功能逐渐受到关注。研究发现,青蛙和小鼠体内有超过10种旁系同源的GTFs(*GTF2A1*基因的转录因子变体)在性腺组织中专一性表达,这些特殊的转录因子不仅承担着所有细胞都必需的基本转录工作,还专门负责调控与繁殖密切相关的关键基因活动[4]。*GTF2A1*基因是卵巢肿瘤检测的一种候选生物标志物,在卵巢癌患者中高度甲基化,提示其可能在生殖系统的功能调控中发挥重要作用[5]。TAO等[6]结合3种基因组方法(GWAS、XP-nSL和ROH)对821只绵羊个体进行全基因组分析,结果将*GTF2A1*基因鉴定为繁殖性状候选基因。狄冉等[7-8]对鲁中肉羊和绵羊的研究发现,*GTF2A1*基因的g.89505005G>A位点多态性与产羔数显著相关,AA基因型个体产羔数显著高于GG基因型个体,这一发现表明,*GTF2A1*基因可能是调控绵羊繁殖性能的关键基因之一。YUE等[2]对澳大利亚白绵羊的研究发现,*GTF2A1*基因的某些插入/缺失(indel)多态性位点与绵羊的产仔数显著相关。LAWAL等[1]采用群体杂合度分析方法对非洲地方鸡品种的正选择信号进行分析,发现

*GTF2A1* 基因与甲状腺刺激激素受体（thyrotropin receptor，TSHR）共同位于一个受选择的基因组区域，该区域 Z（Hp）值最低，而 *TSHR* 基因已被证明在家禽的代谢和繁殖性状中起重要作用[9-10]，提示 *GTF2A1* 基因可能与鸡种对当地气候条件的适应及繁殖性状相关。YUAN 等[11]对白来航鸡和东乡绿壳鸡杂交产生的 $F_2$ 代母鸡进行全基因组关联分析，发现 *GTF2A1* 基因多态性与产蛋数显著相关，推测其可能通过影响卵巢和子宫的功能调控产蛋性能。基于上述研究，可以推测，*GTF2A1* 基因在鸡群体中可能存在影响产蛋性状的功能性遗传变异，有望作为分子标记应用于我国特色鸡种育种实践中。但鸡 *GTF2A1* 基因多态性及其对繁殖性状的调控机制尚未见系统研究报道。苏禽 4 号是江苏省家禽科学研究所自主培育的黄羽肉鸡良种，其母系（M 系）具有繁殖性能稳定、适应性强等特性，是解析基因与性状关联的理想模型。本研究以苏禽 4 号母系为对象，采用 Sanger 测序联合质谱阵列技术（Sequenom MassARRAY）系统筛查其 *GTF2A1* 基因 SNP 位点，通过连锁不平衡分析与单倍型构建，解析 SNP 位点与各产蛋性状的关联性，分析 SNP 位点对保守基序（motif）与转录因子结合能力的影响，旨在筛选优势单倍型组合，为苏禽 4 号母系高繁殖力选育提供分子育种标记。

# 1　材料与方法

## 1.1　试验地点、对象及样品采集

本试验在江苏省家禽科学研究所仪征基地进行。苏禽 4 号母系（M 系）母鸡按照种鸡饲养管理条件采用个体笼养方式饲养。待试验鸡开产后，记录其开产体质量、开产日龄、开产蛋质量、300 日龄蛋质量和 300 日龄产蛋数。在大群母鸡中随机选取 384 只，翅下静脉采血 1 mL/只，加入 0.3 mL 含 5 mg/mL 肝素钠的抗凝剂中，－20 ℃保存。

## 1.2　主要试剂及仪器

*Taq* DNA 聚合酶、dNTPs 及血液/细胞/组织基因组 DNA 提取试剂盒（型号：DP318），均购自天根生化科技（北京）有限公司；紫外分光光度计（NanoDro3 2000）和 Applied Biosystems™（3730xl DNA 分析仪）测序仪，均购自 Thermo Fisher Scientific 公司；MALDI-TOF 质谱仪，购自 Agena Bioscience 公司。

## 1.3　基因组 DNA 提取

采用基因组 DNA 提取试剂盒（血液/细胞/组织）提取鸡全血基因组 DNA，紫外分光光度计（NanoDrop 2000）检测 DNA 浓度与质量（1.8＜$OD_{260}$/ $OD_{280}$＜2.0），稀释至 50 ng/μL，备用。

## 1.4　引物设计与合成

根据 GenBank 上公布的鸡 *GTF2A1* 基因的 DNA 序列（GenBank 登录号：NC_052536.1），用 Primer Primer 5.0 软件在鸡 *GTF2A1* 基因 mRNA 区域设计 6 对 Sanger 测序引物，引物信息见表 1，由上海生工生物技术有限公司合成。针对 Sanger 测序确定的苏禽 4 号母系（M 系）母鸡 5 号染色体基因组 g.40334089 和 g.40341114 位点（版本号：bGal-Gal1.mat.broiler.GRCg7b）序列，使用 sequenom 公司的引物设计软件 Assay design 3.1 设计 MassARRAY SNP 引物组，引物信息见表 2，由金唯智生物科技（北京）有限公司合成。

**表 1　*GTF2A1* 基因 mRNA 区域 Sanger 测序引物相关信息**

Table 1　Information of Sanger sequencing primers for the mRNA regions of the *GTF2A1* gene

| 引物序号 Primer number | 引物序列 Primer sequence | 扩增片段长度/ bp Amplified fragment length | 退火温度/℃ Annealed temperature |
|---|---|---|---|
| P1 | F：5′-GTCACCAGCCTAGGTAAGGA-3′，R：5′-GTTAGTCGAGCTCGCCATGT-3′ | 845 | 60 |
| P2 | F：5′-GGATCCATCATGGCTTCCCTT-3′，R：5′-CTCCTTGCCACACTGGGTAA-3′ | 603 | 58 |
| P3 | F：5′-GTCTGGTGTGTTACCAACCACTA-3′，R：5′-GCTGCATGTCAAGTGCTCTG-3′ | 569 | 55 |
| P4 | F：5′-CATCTCAAGGATGGCATCATGA-3′，R：5′-CCAAGGACTCAGGGAAAGA-3′ | 1 542 | 65 |
| P5 | F：5′-GAAGGACTTGCAGAAGCAGACA-3′，R：5′-CCATCAGTCATCACGGTTACA-3′ | 1 384 | 63 |
| P6 | F：5′-CCTGAGTTGCTGGTTAGTGACA-3′，R：5′-GAGGAGCAACCTTGATCTGCAA-3′ | 1 087 | 62 |

## 1.5　鸡 *GTF2A1* 基因 SNP 位点确定

将 384 份母鸡个体 DNA 样品各取 1 μL，置于 5 mL 灭菌离心管中混合，制备 DNA 混合池模板。使用 6 对 Sanger 测序引物（表 1）进行 PCR 扩增，反应体系为：10×PCR 缓冲液（含 15 mmol/L $Mg^{2+}$）2.5 μL，10 μmol/L 上下游引物各 1.0 μL，2 mmol/L dNTPs 2.0 μL，5.0 U/μL *Taq* DNA 聚合酶 0.5 μL，50 ng/μL pooled DNA 模板 1.0 μL，$ddH_2O$ 补足

25 μL。PCR扩增程序设定为:94 ℃预变性5 min;94 ℃变性30 s,58~65 ℃(具体依表1中的引物而定)退火15 s,72 ℃延伸30 s,35个循环;72 ℃延伸10 min;最后4 ℃保存。PCR扩增产物经1.5%琼脂糖凝胶电泳检测后,委托生工生物工程(上海)股份有限公司在ABI 3730xl DNA分析仪上进行Sanger测序。利用DNAstar软件对获得的序列进行分析,筛选SNP位点。利用欧洲生物信息学研究所(EBI)旗下的Ensembl数据库(http://asia.ensembl.org/index.html?redirect=no),基于鸡的参考基因组(版本号:bGalGal1.mat.broiler.GRCg7b),对*GTF2A1*基因的SNP位点进行全基因组定位。

**表2 *GTF2A1*基因mRNA区域SNP位点引物相关信息**

Table 2 Information of primers of SNP loci in the mRNA regions of the *GTF2A1* gene

| SNP位点 SNP locus | SNP物理位置 SNP physical position | PCR扩增引物 Amplification primer | | | 单碱基延伸引物 Single base extension primer | |
|---|---|---|---|---|---|---|
| | | 序列 Sequence | 产物长度/bp Length of product | 产物位置 Location of product | 方向 Direction | 序列 Sequence |
| g.40334089G>C | rs40334089 | F:5′-ACGTTGGATGCTCTAAAGTTTCCAGGTGGC-3′<br>R:5′-ACGTTGGATGCCAAGTCAGTCAACAGACAC-3′ | 113 | 5′非翻译区 5′UTR | F | GCTGCACAGCTTTGAGAA |
| g.40341114A>G | rs40341114 | F:5′-ACGTTGGATGACCAGCTCAAGCACAGATTC-3′<br>R:5′-ACGTTGGATGGCTCCATCAACTTGGAGAAC-3′ | 116 | 外显子3 Exon 3 | R | tcGGTTGTGCCTGTTGTTGCTG |

### 1.6 鸡*GTF2A1*基因Sequenom MassARRAY-SNP位点确定

采用多重PCR技术(引物见表2)同步扩增包含多个目标位点的DNA片段。按比例混合以下组分制备母液:HPLC级水691 μL,10×PCR Buffer 236 μL,25 mmol/L $MgCl_2$ 123 μL,25 mmol/L dNTPs 38 μL,0.5 μmol/L上下游引物混合物378 μL,5 U/μL *Taq* DNA聚合酶76 μL。母液充分混匀,分装至384孔PCR板,每孔4 μL。随后每孔加入10 ng/μL基因DNA模板1 μL,使单孔反应液体积为5 μL。多重PCR扩增程序:94 ℃预变性2 min;94 ℃变性20 s,56 ℃退火30 s,72 ℃延伸60 s,循环45次;72 ℃延伸3 min;4 ℃保存。

多重反应产物用虾碱性磷酸酶(shrimp alkaline phosphatase,SAP)进行消化处理,降解残留dNTPs并灭活游离引物,以纯化多重PCR扩增产物。SAP消化反应体系:将810.9 μL HPLC级水、90.1 μL 10×SAP缓冲液、159 μL 1.7 U/μL SAP混合配制成母液,按24 μL/孔分装至384孔板;反应程序为37 ℃温育40 min后85 ℃热失活5 min。在*iPLEX*酶催化下,使用针对各靶位点设计的单碱基延伸引物(表2),以双脱氧核苷酸(ddNTPs)为终止底物,对纯化后的多重反应产物进行单碱基延伸反应,延伸反应体系:将400.2 μL HPLC级水、106 μL 10×iPLEX Buffer Plus、106 μL iPLEX Termination Mix、426.1 μL 1.0 μmol/L延伸引物混合物及22 μL *iPLEX*酶混合配制成母液,按2 μL/孔分装至384孔板。延伸反应程序为:94 ℃预变性30 s;94 ℃变性5 s,52 ℃退火5 s,80 ℃延伸5 s,退火延伸5个循环,变性退火延伸40个循环;72 ℃延伸3 min;4 ℃保存。

对延伸产物进行3倍稀释后,采用离子交换树脂进行脱盐纯化,点样至384孔SpectroCHIP芯片。利用基质辅助激光解吸电离飞行时间质谱(MALDI-TOF MS)平台(Sequenom MassARRAY系统)检测。不同等位基因由于延伸引物结合ddNTPs的分子量差异,产生特异性质量位移峰。最终采用Typer 4.0软件解析质谱峰图谱,实现SNP位点的自动化基因分型。

### 1.7 统计分析

1.7.1 *遗传数据分析* 基于苏禽4号母系(M系)母鸡群体,利用Excel软件计算*GTF2A1*基因g.40334089G>C与g.40341114A>G位点的等位基因频率(allele frequency,AF)、基因型频率(genotype frequency,GF)、期望杂合度(expected heterozygosity,He)、多态信息含量(polymorphism information content,PIC)及有效等位基因数(effective number of alleles,Ne);通过卡方检验($\chi^2$)评估Hardy-Weinberg平衡状态,并采用Hapoview v4.1软件分析2个SNP位点间的连锁不平衡(linkage disequilibrium,LD)。

1.7.2 *SNP位点与产蛋性状的关联分析* 采用IBM SPSS Statistics 27.0软件,对*GTF2A1*基因SNP位点与开产日龄(age at first egg,AFE)、开产

体质量（body weight at first egg，BWF）、开产蛋质量（weight of the first egg，WFE）、300 日龄平均蛋质量（average egg weight at 300 days of age，EW300）及 300 日龄产蛋数（egg number to 300 days of age，EN300）的关联性进行单变量方差分析（UNIANOVA），评估基因型对产蛋性状的影响。数据以“平均值±标准误”表示，基因型间差异显著性经 LSD 法多重比较判定，$P<0.05$ 为差异显著。

1.7.3　转录因子结合位点分析　利用 MEME 网站（https://meme-suite.org）分析 *GTF2A1* 基因 SNP 位点短序列模式保守基序[12]，利用 Tomtom 插件对 *GTF2A1* 基因 SNP 位点保守基序转录因子结合区域进行分析[13]，利用 GOMo 插件对 *GTF2A1* 基因 SNP 位点保守基序进行功能预测[14]。

## 2　结果与分析

### 2.1　鸡 *GTF2A1* 基因 SNP 位点筛选

对引物 P1 至 P6 的 DNA 混合池样本的扩增产物进行测序，在苏禽 4 号母系（M 系）母鸡 *GTF2A1* 基因 5′非翻译区（5′ UTR）区鉴定到 1 个 SNP 位点（g.40334089G>C），在外显子 3（exon 3）区鉴定到 1 个 SNP 位点（g.40341114A>G）（图 1）。经 Ensembl 数据库分析发现，*GTF2A1* 基因转录起始是反向的，因此 g.40334089G>C 突变未引起调控元件的变化，为颠换突变；g.40341114A>G 突变属于转换突变，未引发氨基酸改变，为同义突变。

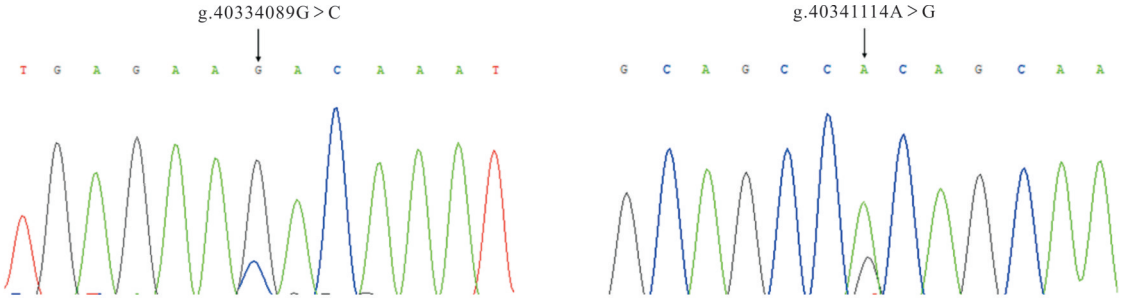
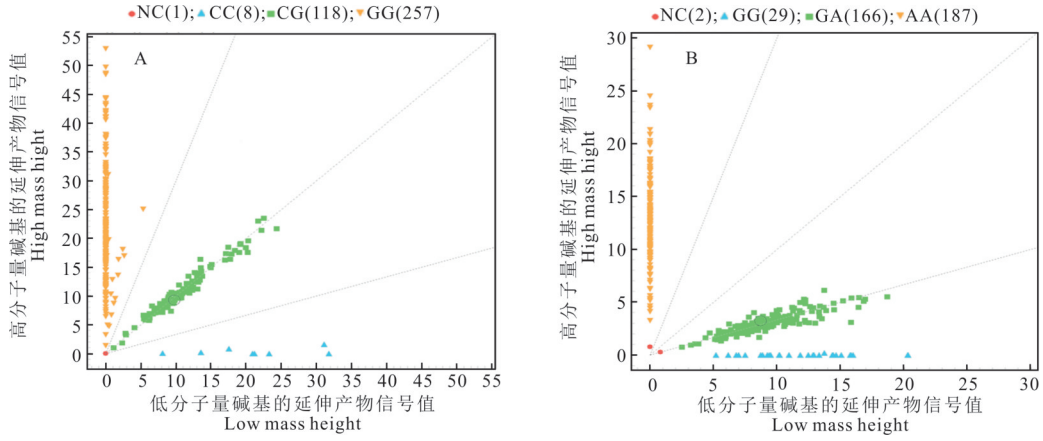



图 1　鸡 *GTF2A1* 基因 2 个 SNP 位点 Sanger 测序图

Fig. 1　Sanger sequencing map of two SNPs in the chicken *GTF2A1* gene

### 2.2　鸡 *GTF2A1* 基因 SNP 位点质谱分析

384 只苏禽 4 号母系（M 系）母鸡 *GTF2A1* 基因 2 个 SNP 位点基因型质谱分析结果见图 2。




A. g.40334089G>C 位点质谱分型结果，NC（1）表示无信号个体数为 1，CC（8）、CG（118）、GG（257）分别表示 CC、CG 和 GG 基因型个体数为 8，118 和 257 只；B. g.40341114A>G 位点质谱分型结果，NC（2）表示无信号个体数为 2，GG（29），GA（166），AA（187）分别表示 GG、GA 和 AA 基因型个体数为 29，116 和 187 只。

A. MassARRAY genotyping results for the g.40334089G>C locus. NC（1）indicates one individual with no signal；CC（8），CG（118），and GG（257）indicate that the numbers of individuals with the CC，CG，and GG genotypes are 8，118，and 257，respectively. B. MassARRAY genotyping results for the g.40341114A>G locus. NC（2）indicates two individuals with no signal；GG（29），GA（166），and AA（187）indicate that the numbers of individuals with the GG，GA，and AA genotypes are 29，166，and 187，respectively.

图 2　鸡 *GTF2A1* 基因 2 个 SNP 位点质谱分型图

Fig. 2　Genotyping spectrograms obtained by MassARRAY for two SNP loci in the chicken *GTF2A1* gene

由图 2 可知，苏禽 4 号母系（M 系）母鸡 *GTF2A1* 基因 g.40334089G>C 位点存在 GG、CG

和CC等3种基因型(图2-A),检出率为99.7%(383/384);g.40341114A>G位点存在AA、GA和GG等3种基因型(图2-B),检出率为99.5%(382/384)。

### 2.3 鸡*GTF2A1*基因SNP位点群体遗传学分析

分别根据*GTF2A1*基因g.40334089G>C和g.40341114A>G位点分型结果,统计苏禽4号母系(M系)母鸡2个位点的基因型频率和等位基因频率,结果见表3。由表3可知,苏禽4号母系(M系)在g.40334089G>C位点以GG基因型为主,G等位基因频率明显高于C等位基因;g.40341114A>G位点以AA基因型为主,A等位基因频率明显高于G等位基因。Hardy-Weinberg平衡检验结果(表4)显示,2个SNP位点均处于Hardy-Weinberg平衡状态,g.40334089G>C位点He和PIC分别为0.29和0.25,g.40341114A>G位点He和PIC分别为0.42和0.33。鉴于g.40334089G>C和g.40341114A>G位点的期望杂合度数值均落在0.25~0.5,且多态信息含量均未超过0.5的中度多态阈值,据此判定2个SNP位点均为中度多态。

**表3 鸡*GTF2A1*基因2个SNP位点等位基因频率和基因型频率**

Table 3 Allele and genotype frequencies of two SNP loci in the chicken *GTF2A1* gene

| SNP位点<br>SNP locus | 基因型<br>Genotype | 个体数<br>No. of samples | 基因型频率<br>Genotype frequency | 等位基因<br>Allele | 等位基因频率<br>Allele frequency |
|---|---|---|---|---|---|
| g.40334089G>C | CC | 8 | 0.02 | C | 0.17 |
| | CG | 118 | 0.31 | G | 0.83 |
| | GG | 257 | 0.67 | | |
| g.40341114A>G | AA | 187 | 0.49 | A | 0.71 |
| | GA | 166 | 0.43 | G | 0.29 |
| | GG | 29 | 0.08 | | |

**表4 鸡*GTF2A1*基因2个SNP位点的Hardy-Weinberg平衡检测与群体遗传学参数**

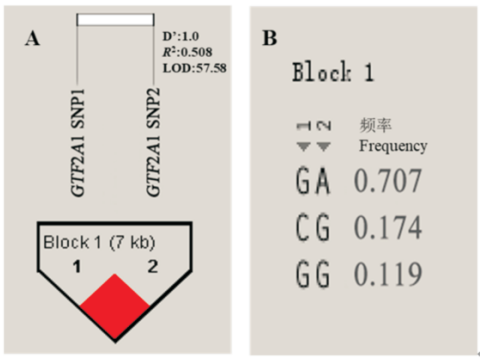
Table 4 Hardy-Weinberg equilibrium testing and population genetic parameters for two SNP loci in the chicken *GTF2A1* gene

| SNP位点<br>SNP locus | Hardy-Weinberg平衡检测<br>Hardy-Weinberg equilibrium testing | | 期望杂合度<br>Heterozygosity | 杂合度方差<br>Heterozygosity variance | 多态信息含量<br>Polymorphism information content | 有效等位基因数<br>Effective number of alleles |
|---|---|---|---|---|---|---|
| | $\chi^2$ | $P$ | | | | |
| g.40334089G>C | 1.74 | >0.05 | 0.29 | 0.05 | 0.25 | 1.41 |
| g.40341114A>G | 0.90 | >0.05 | 0.42 | 0.06 | 0.33 | 1.71 |

### 2.4 鸡*GTF2A1*基因2个SNP位点连锁不平衡与单倍型分析

LD分析结果见图3。由图3-A可知,本研究*GTF2A1*基因g.40334089G>C和g.40341114A>G位点的LOD值为57.58,依据MORTON[15]提出的连锁分析显著性标准(LOD>3.0,对应$P<0.0001$)可知,这2个SNP位点存在连锁的概率极高,显著降低了假阳性风险。*GTF2A1*基因g.40334089G>C和g.40341114A>G位点的$|D'|$值为1.0,$R^2$值为0.508,依据经典阈值标准[16]($|D'|>0.80$且$R^2>0.33$)可知,2个SNP位点处于高强度连锁不平衡状态。

g.40334089G>C与g.40341114A>G之间的强连锁关系产生了3种单倍型:H1(GA,频率=0.707)、H2(CG,频率=0.174)和H3(GG,频率=0.119)(图3-B)。单倍型组合形成的基因型及其频率详见表5。由表5可知,本研究共检测到6种单倍型组合个体,其中H1H1和H2H1为优势单倍型组合;未检测到H1H2、H1H3和H3H2单倍型组合个体。




A. 连锁图,SNP1代表g.40334089G>C位点,SNP2代表g.40341114A>G位点;B. 单倍型分析。

A. Linkage map, where SNP1 represents the g.40334089G>C locus and SNP2 represents the g.40341114A>G locus; B. Haplotype analysis.

图3 鸡*GTF2A1*基因2个SNP位点连锁不平衡分析

Fig. 3 Linkage disequilibrium analysis of two SNP loci in the chicken *GTF2A1* gene

**表 5　鸡 *GTF2A1* 基因 2 个 SNP 位点单倍型组合、基因型和基因型频率**

Table 5　Haplotype combinations, genotypes, and genotype frequencies of two SNP loci in the chicken *GTF2A1* gene

| 单倍型组合 Haplotype combination | 基因型 Binding genotype | 个体数 No. of samples | 基因型频率 Genotype frequency | 单倍型组合 Haplotype combination | 基因型 Binding genotype | 个体数 No. of samples | 基因型频率 Genotype frequency |
|---|---|---|---|---|---|---|---|
| H1H1 | GGAA | 187 | 0.49 | H2H3 | CGGG | 17 | 0.04 |
| H2H1 | CGGA | 100 | 0.26 | H3H1 | GGGA | 66 | 0.17 |
| H2H2 | CCGG | 8 | 0.02 | H3H3 | GGGG | 4 | 0.01 |

### 2.5　鸡 *GTF2A1* 基因 SNP 位点多态性与产蛋性状的关联性分析

对苏禽 4 号母系（M 系）母鸡群体 *GTF2A1* 基因 g.40334089G>C 和 g.40341114A>G 位点与产蛋性状进行关联性分析，结果见表 6。由表 6 可知，g.40334089G>C 和 g.40341114A>G 位点不同基因型之间的部分产蛋性状指标存在显著差异，其中 g.40334089G>C 位点 GG 基因型个体开产日龄显著早于 CC 及 CG 基因型个体（$P<0.05$），300 日龄产蛋数显著高于 CC 和 CG 基因型个体（$P<0.05$）；CC 基因型个体 300 日龄蛋质量显著高于 CG 与 GG 基因型个体（$P<0.05$），GG 基因型个体的 300 日龄蛋质量最低；g.40341114A>G 位点 AA 基因型个体开产日龄显著早于 GA 与 GG 基因型个体（$P<0.05$），300 日龄产蛋数显著高于 GA 和 GG 基因型个体（$P<0.05$）。

**表 6　鸡 *GTF2A1* 基因 2 个 SNP 位点不同基因型与产蛋性状的关联分析**

Table 6　Association analysis between different genotypes of two SNP loci in the chicken *GTF2A1* gene and egg-laying traits

| 多态位点 Polymorphic site | 基因型 Genotype | 个体数 No. of samples | 开产日龄/d Age at first egg | 开产体质量/（g·只$^{-1}$） Body weight at first egg | 开产蛋质量/（g·枚$^{-1}$） First egg weight | 300 日龄蛋质量/（g·枚$^{-1}$） Egg weight at 300 days | 300 日龄产蛋数/枚 Total egg production at 300 days |
|---|---|---|---|---|---|---|---|
| g.40334089G>C | CC | 8 | 158.13±5.51 b | 2 044.13±328.33 | 34.96±4.13 | 51.50±5.15 a | 88.00±11.98 b |
| | CG | 118 | 156.19±11.26 b | 2 104.98±283.52 | 36.92±4.18 | 49.25±3.52 b | 87.43±13.42 b |
| | GG | 257 | 149.10±10.10 a | 2 103.73±289.48 | 36.88±4.16 | 48.94±3.70 b | 96.05±10.28 a |
| g.40341114A>G | AA | 187 | 148.15±9.66 a | 2 096.51±297.89 | 37.12±4.07 | 48.81±3.64 | 98.62±8.42 a |
| | GA | 166 | 154.85±11.22 b | 2 121.39±265.57 | 36.66±4.35 | 49.26±3.80 | 87.87±12.51 b |
| | GG | 29 | 153.38±10.63 b | 2 037.83±339.68 | 36.27±3.60 | 49.92±3.29 | 89.28±14.02 b |

注：同一位点同列数据后标不同小写字母表示差异显著（$P<0.05$）。

Note: Different lowercase letters following data within the same column for a given locus indicate significant differences ($P<0.05$).

### 2.6　鸡 *GTF2A1* 基因单倍型组合与产蛋性状的关联性分析

对苏禽 4 号母系（M 系）群体中 *GTF2A1* 基因单倍型组合与产蛋性状的关联性进行分析，结果如表 7 所示。

**表 7　鸡 *GTF2A1* 基因单倍型组合与产蛋性状的关联分析**

Table 7　Association analysis between haplotype combinations of the chicken *GTF2A1* gene and egg-laying traits

| 单位型组合 Haplotype combination | 基因型 Binding genotype | 个体数 No. of samples | 开产日龄/d Age at first egg | 开产体质量/（g·只$^{-1}$） Body weight at first egg | 开产蛋质量/（g·枚$^{-1}$） First egg weight | 300 日龄蛋质量/（g·枚$^{-1}$） Egg weight at 300 days | 300 日龄产蛋数/枚 Total egg production at 300 days |
|---|---|---|---|---|---|---|---|
| H1H1 | GGAA | 187 | 148.15±9.66 bc | 2 096.51±297.89 | 37.13±4.07 | 48.82±3.64 a | 98.62±8.42 a |
| H2H1 | CGGA | 100 | 156.76±11.14 ab | 2 111.53±269.96 | 36.91±4.33 | 49.29±3.72 | 87.09±12.90 b |
| H2H2 | CCGG | 8 | 158.13±5.51 a | 2 044.13±328.33 | 34.96±4.13 | 51.50±5.15 b | 88.00±11.98 b |
| H2H3 | CGGG | 17 | 152.84±11.68 abc | 2 066.47±360.58 | 37.00±3.27 | 49.06±2.01 | 89.41±16.46 b |
| H3H1 | GGGA | 66 | 151.96±10.80 abc | 2 136.32±260.11 | 36.27±4.39 | 49.20±3.93 | 89.05±11.89 b |
| H3H3 | GGGG | 4 | 146.17±11.12 c | 1 903.50±321.12 | 35.75±4.04 | 50.42±2.54 | 91.25±6.75 ab |

注：同列数据后标不同小写字母表示差异显著（$P<0.05$）。

Note: Different lowercase letters following data within the same column indicate significant differences ($P<0.05$).

表 7 显示，不同单倍型组合个体 300 日龄产蛋数存在显著差异（$P<0.05$），其中 H1H1（GGAA）单倍型组合个体的 300 日龄产蛋数最高，显著高于 H2H1（CGGA）、H2H2（CCGG）、H2H3（CGGG）及 H3H1（GGGA）单倍型组合个体（$P<0.05$）；与 H3H3（GGGG）单倍型组合个体的 300 日龄产蛋数无显著差异，但比其多 7.37 枚。不同单倍型组合个体的开产日龄存在显著差异，单倍型组合 H3H3

(GGGG)个体的开产日龄显著早于单倍型组合H2H1(CGGA)和H2H2(CCGG)个体($P<0.05$),与单倍型组合H1H1(GGAA)、H2H3(CGGG)和H3H1(GGGA)个体的开产日龄差异虽不显著,但比其分别早1.98,6.67和5.79 d。

### 2.7 鸡 *GTF2A1* 基因 SNP 位点碱基变化对保守基序和转录因子的影响

利用MEME网站分析*GTF2A1*基因突变位点的保守基序,结果发现g.40341114A>G位点有一个保守基序:CCTGCAGCTGGTCAGCAGCAGCC-R(A/G)CAGC,g.40334089G>C位点没有保守基序。用Tomtom分析*GTF2A1*基因SNP位点保守基序与转录因子的结合能力,发现g.40341114A>G位点A→G突变使保守基序与转录因子Myod1的结合能力增强(图4-A);G→A突变使保守基序与转录因子REST和PPARG的结合能力增强(图4-B,C),用GOMO对*GTF2A1*基因g.40341114A>G位点保守基序进行功能预测,结果发现该保守基序参与了小GTP酶介导的信号转导、蛋白质转运、葡萄糖代谢过程等生物过程及趋化因子活性、GTP结合等分子功能(图5)。

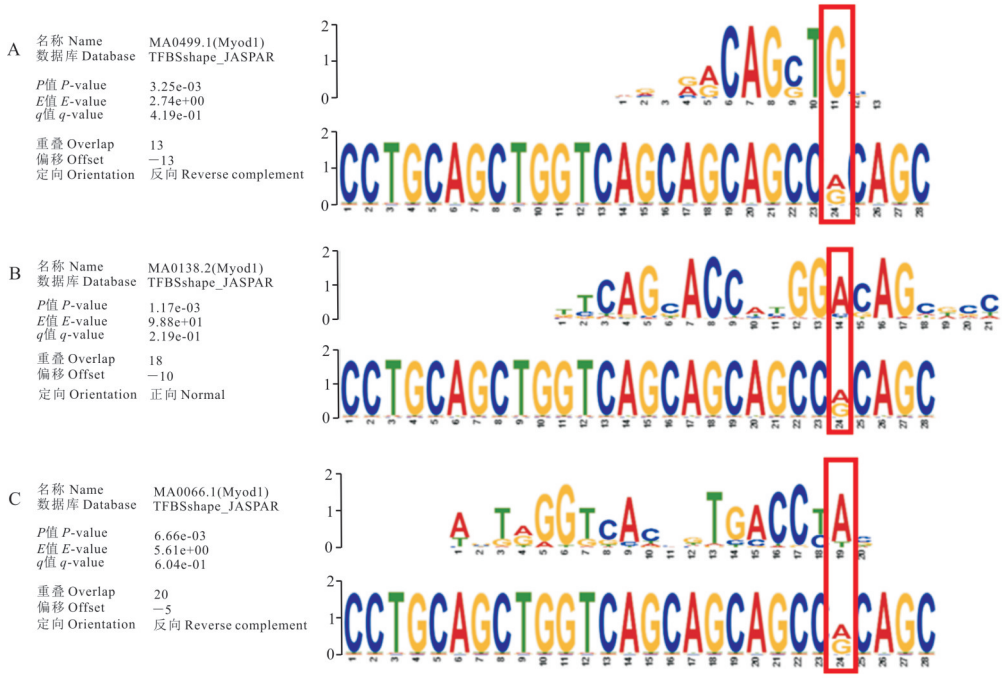



图4 鸡*GTF2A1*基因 g.40341114A>G 位点保守基序分析

Fig. 4 Analysis of the conserved motif at the g.40341114A>G locus in the chicken *GTF2A1* gene




图5 鸡*GTF2A1*基因 g.40341114A>G 位点保守基序功能预测

Fig. 5 Functional annotation of the conserved motifs at the g.40341114A>G locus in the chicken *GTF2A1* gene

## 3 讨 论

### 3.1 鸡 *GTF2A1* 基因群体遗传学分析

本试验对苏禽4号母系(M系)mRNA区域进行一代测序分析,结果发现了2个SNP位点,即g.40334089G>C和g.40341114A>G位点,其中G→C属于颠换突变,A→G属于转换突变,且为同义突变。这2个SNP位点均有3种基因型,均处于Hardy-Weinberg平衡状态。遗传参数分析结果表明,在所检测的2个中性位点上,苏禽4号母系(M系)群体目前没有受到强烈的近亲繁殖[17]、选择或其他导致基因型频率失衡因素[18]的影响,符合随机交配大群体的基本特征。该结果为后续的关联分析等研究提供了合格的群体遗传学基础。群体的遗传丰富度与遗传多样性呈正相关,遗传多样性常通过期望杂合度和多态信息含量等指标进行度

量[19]。g. 40334089G>C 和 g. 40341114A>G 位点均为中度多态，表明本试验的育种群拥有丰富的遗传变异可供选择和利用，世代扩繁的后代具有更高的遗传多样性，能产生更强的杂种优势，且具有更强的环境变化适应能力，可减小近交衰退的风险，具有更大的选择潜力。

### 3.2　鸡 *GTF2A1* 基因多态性与产蛋性状的关联性分析

*GTF2A1* 基因被鉴定为现代家鸡中与高产蛋频率相关的驯化基因[9-10,20]。YUAN 等[11]用 GWAS 方法对白来航鸡和东乡绿壳蛋鸡正反交后代进行全基因组扫描，共鉴定出 3 个与产蛋性状相关的 SNP 位点，其中位于 5 号染色体 *GTF2A1* 基因上的 rs317449530（*Gallus gallus* 4.0 版本基因组位置：40101576，3′ UTR）A→G 突变和 rs313187645（*Gallus gallus* 4.0 版本基因组位置：40106943，内含子 7）A→G 突变均与产蛋前期（21 至 26 周龄）产蛋数显著相关，rs317449530、rs313187645 和 rs317282705（*Gallus gallus* 4.0 版本基因组位置：40121093，5′UTR）C→T 突变均与 21 至 40 周龄产蛋数显著相关。本试验发现，g. 40334089G>C 和 g. 40341114A>G 位点均与苏禽 4 号母系（M 系）的开产日龄和 300 日龄产蛋数显著相关。本试验与 YUAN 等[11]的研究结果在 SNP 位点和相关性分析结果方面存在差异，这可能是由品种及参考基因组不同引起的。

基因突变可发生在基因组多个功能区域，包括非翻译区（UTR）、编码区和内含子区域，不同位点的突变通过不同分子机制影响基因表达、蛋白质功能或调控网络，进而导致表型变异[21]。本试验中 g. 40334089G>C 位于 5′ UTR，g. 40341114A>G 位于外显子 3 区域，g. 40334089G>C 没有引起调控元件的变化，g. 40341114A>G 为同义突变，但 5′ UTR 突变可干扰 mRNA 翻译起始复合物的组装[22]，外显子区域的突变也可能引起蛋白质合成异常[23]，从而导致这 2 个位点突变均与开产日龄和 300 日龄产蛋数显著相关。

蛋白质-蛋白质相互作用通常由一个蛋白质中的保守基序和另一个蛋白质中的结构域介导，从而引起细胞功能的变化[24]。本试验 g. 40341114A>G 位点 G→A 突变使保守基序与转录因子 REST 和 PPARG 的结合能力增强，该位点 AA 基因型个体的开产日龄更早，300 日龄产蛋数更高，说明保守基序与 REST 和 PPARG 结合能力的增强可能调控了 *GTF2A1* 基因的表达，从而使产蛋性状得到提升。

表型变异由单一位点突变或多 SNP 连锁效应介导[25]。通过连锁不平衡分析定位共遗传位点，可进一步揭示非等位基因的多位点互作，从而系统解析表型变异的遗传基础[26]。本研究发现，*GTF2A1* 基因 g. 40334089G>C 和 g. 40341114A>G 位点存在强连锁不平衡。交配模式、群体混合与选择压力共同调控群体内 SNP 连锁不平衡的构建与维持[27]。本试验选用的苏禽 4 号母系（M 系）经过 6 个世代的选育，选择压力较强，可能导致了 2 个 SNP 位点的强连锁。2 个位点的强连锁出现 3 种单倍型，其中 H1（GA）单倍型频率最高（0.707），H3（GG）单倍型频率最低（0.119），根据 AKEY[28]理论，高频率 GA 可能为祖先中最早出现的单倍型。本研究中仅检测到 6 种单倍型组合，少于理论值（9 种），其原因可能是样本量有限未能覆盖低频组合，或自然选择对某些单倍型组合产生了淘汰作用。

本试验 g. 40334089G>C 和 g. 40341114A>G 2 个 SNP 位点强连锁产生的单倍型组合 H1H1（GGAA）的基因型频率最高（0.49），且与 300 日龄产蛋数显著相关（$P<0.05$），个体 300 日龄产蛋数最高，平均为 98.62 枚，显著高于其他单倍型组合个体；单倍型组合 H3H3（GGGG）与开产日龄显著相关（$P<0.05$），其个体开产日龄最早，为 146.17 d，但其基因型频率较低（0.01）。本试验发现，GGAA 基因型个体 300 日龄产蛋数虽然最高，但开产日龄比 GGGG 基因型个体迟 1.98 d，GGGG 基因型个体的开产日龄虽然最早，但 300 日龄产蛋数比 GGAA 基因型个体少 7.37 枚，因此这 2 个基因型在产蛋性状上表现出权衡关系。种鸡选育以提升总产蛋数为首要目标，因此 GGAA 基因型具有潜在优势。单一位点多态性分析对揭示基因-性状相关性存在局限性，同时对多个位点的综合效应进行分析才能准确得出基因与性状的真实相关性[29]。本研究提示，在苏禽 4 号母系（M 系）育种中可将单倍型组合 H1H1（GGAA）作为产蛋性状的重要分子标记，优先选择 GGAA 基因型个体作为亲本，以加速母系选育进程。本试验结果为产蛋性状的分子标记辅助选择提供了依据。

## 4　结　论

在苏禽 4 号母系（M 系）*GTF2A1* 基因 mRNA 区筛选到 g. 40334089G>C 和 g. 40341114A>G 位

点,2个SNP位点均存在3种基因型且符合Hardy-Weinberg平衡,均与苏禽4号母系(M系)母鸡的开产日龄及300日龄产蛋数显著相关。2位点强连锁,形成H1(GA)、H2(CG)、H3(GG)3种单倍型,其中优势单倍型H1(GA)频率为0.707。H1H1(GGAA)单倍型组合个体的300日龄产蛋数显著高于除H3H3(GGGG)单倍型组合个体外的其他组合($P<0.05$);H3H3(GGGG)单倍型组合个体开产日龄最早;H1H1(GGAA)单倍型组合个体在开产日龄与300日龄产蛋数上综合效应最优。单倍型组合H1H1(GGAA)可作为苏禽4号母系(M系)产蛋性状分子标记辅助育种的候选标记。

## [参考文献]


[1] LAWAL R A, AL-ATIYAT R M, ALJUMAAH R S, et al. Whole-genome resequencing of red junglefowl and indigenous village chicken reveal new insights on the genome dynamics of the species[J]. Frontiers in Genetics, 2018, 9: 264.

[2] YUE L X, LI W J, PAN Y J, et al. Polymorphism of GTF2A1 gene is associated with litter size in sheep[J]. Theriogenology, 2023, 208: 194-200.

[3] KOBAYASHI N, BOYER T G, BERK A J. A class of activation domains interacts directly with TFIIA and stimulates TFIIA-TFIID-promoter complex assembly[J]. Molecular and Cellular Biology, 1995, 15(11): 6465-6473.

[4] FREIMAN R N. Specific variants of general transcription factors regulate germ cell development in diverse organisms[J]. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 2009, 1789(3): 161-166.

[5] HUANG Y W, JANSEN R A, FABBRI E, et al. Identification of candidate epigenetic biomarkers for ovarian cancer detection[J]. Oncology Reports, 2009, 22(4): 853-861.

[6] TAO L, WANG X Y, ZHONG Y J, et al. Combined approaches identify known and novel genes associated with sheep litter size and non-seasonal breeding[J]. Animal Genetics, 2021, 52(6): 857-867.

[7] 狄冉,刘秋月,王翔宇,等. 绵羊*GTF2A1*基因多态性及其与产羔数的关联分析[J]. 中国畜牧杂志,2019,55(11):63-67.
DI R, LIU Q Y, WANG X Y, et al. Polymorphism of sheep *GTF2A1* gene and its association with lamb rate[J]. Chinese Journal of Animal Science, 2019, 55(11): 63-67.

[8] 狄冉,刘秋月,潘林香,等. 鲁中肉羊*GTF2A1*基因多态性及其与产羔数的关联分析[J]. 中国草食动物科学,2019,39(4):15-17.
DI R, LIU Q Y, PAN L X, et al. Polymorphism of *GTF2A1* gene and its association with litter size in Luzhong mutton sheep[J]. China Herbivore Science, 2019, 39(4): 15-17.

[9] YOSHIMURA T, YASUO S, WATANABE M, et al. Light-induced hormone conversion of T4 to T3 regulates photoperiodic response of gonads in birds[J]. Nature, 2003, 426(6963): 178-181.

[10] RUBIN C J, ZODY M C, ERIKSSON J, et al. Whole-genome resequencing reveals loci under selection during chicken domestication[J]. Nature, 2010, 464(7288): 587-591.

[11] YUAN J W, SUN C J, DOU T C, et al. Identification of promising mutants associated with egg production traits revealed by genome-wide association study[J]. PLoS One, 2015, 10(10): e0140615.

[12] BAILEY T L, ELKAN C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers[J]. Proceedings International Conference on Intelligent Systems for Molecular Biology, 1994, 2: 28-36.

[13] GUPTA S, STAMATOYANNOPOULOS J A, BAILEY T L, et al. Quantifying similarity between motifs[J]. Genome Biology, 2007, 8(2): R24.

[14] BUSKE F A, BODÉN M, BAUER D C, et al. Assigning roles to DNA regulatory motifs using comparative genomics[J]. Bioinformatics, 2010, 26(7): 860-866.

[15] MORTON N E. Sequential tests for the detection of linkage[J]. American Journal of Human Genetics, 1955, 7(3): 277-318.

[16] ARDLIE K G, KRUGLYAK L, SEIELSTAD M. Patterns of linkage disequilibrium in the human genome[J]. Nature Reviews Genetics, 2002, 3(4): 299-309.

[17] HARTL D L, CLARK A G. Principles of population genetics[M]. 4th ed. Sunderland, MA: Sinauer Associates, 2007: 137-138.

[18] HEDRICK P W. Genetics of populations[M]. Boston: Science Books International, 1983: 232-234.

[19] ZANETTI E, DE MARCHI M, ABBADI M, et al. Variation of genetic diversity over time in local Italian chicken breeds undergoing in situ conservation[J]. Poultry Science, 2011, 90(10): 2195-2201.

[20] HANON E A, LINCOLN G A, FUSTIN J M, et al. Ancestral TSH mechanism signals summer in a photoperiodic mammal[J]. Current Biology, 2008, 18(15): 1147-1152.

[21] SCHELL R, HALE J J, MULLIS M N, et al. Genetic basis of a spontaneous mutation's expressivity[J]. Genetics, 2022, 220(3): iyac013.

[22] GRAYESKI P J, WEIDMANN C A, KUMAR J, et al. Global 5′-UTR RNA structure regulates translation of a SERPINA1 mRNA[J]. Nucleic Acids Research, 2022, 50(17): 9689-9704.

[23] CARTEGNI L, CHEW S L, KRAINER A R. Listening to silence and understanding nonsense: exonic mutations that affect splicing[J]. Nature Reviews Genetics, 2002, 3(4): 285-298.

[24] IDREES S, PAUDEL K R, SADAF T, et al. Uncovering domain motif interactions using high-throughput protein-protein interaction detection methods[J]. FEBS Letters, 2024, 598(7): 725-742.


(下转第38页)