

# 全基因组关联分析(GWAS)鉴别猪腹股沟/阴囊疝易感基因

阮国荣<sup>2</sup>, 龙毅<sup>1</sup>, 宿英<sup>1</sup>, 张志燕<sup>1</sup>, 肖石军<sup>1</sup>, 廖信军<sup>1,3</sup>, 艾华水<sup>1</sup>, 杨斌<sup>1</sup>,  
邓政<sup>1</sup>, 辛文水<sup>1</sup>, 唐建红<sup>1</sup>, 任军<sup>1</sup>, 丁能水<sup>1\*</sup>, 黄路生<sup>1</sup>

(1. 江西农业大学 动物生物技术国家重点实验室培育基地, 南昌 330045; 2. 福建农业职业技术学院, 福州 350119;  
3. 井冈山大学生命科学学院, 吉安 343009)

**摘要:** 旨在鉴别猪腹股沟/阴囊疝的易感基因并探寻其发病机理, 本研究对源于纯种长白和大白 2 个西方商业猪种 220 个腹股沟/阴囊疝三联体核心家系群(包含 237 个患病个体)的 734 个样本, 利用 Illumina 猪 60K SNP 芯片进行扫描和基因分型, 通过基于 SNP 的病例-对照全基因组关联分析, 结果发现, 在 SSC1 上 19.31~19.59 Mb 区域内的 4 个 SNPs 标记以及 SSC7 上 17.05 Mb 处有 1 个 SNP 标记与猪腹股沟/阴囊疝呈显著相关。其中, SSC1 上的易感区域内存在 1 个位置候选基因编码糖醛基-2-磺基转移酶(*UST*), 该酶所调控的硫酸皮肤素粘多糖, 其代谢紊乱可引起人类伴发各种疝气的粘多糖症; 而 SSC7 上易感位点位于编码细胞周期蛋白依赖激酶 5 调节亚单位相关蛋白 1 类似物 1 的 *CDKAL1* 基因内, 推测该基因可能通过调控细胞凋亡影响疝气发生。但在扩大群体至 1 134 个个体的(包含 367 个腹股沟/阴囊疝患病个体)后仅位于 *CDKAL1* 基因中的 MARC0077275 位点在 TDT 分析中得到重复验证。本研究通过全基因组关联分析和大规模家系群体中的 TDT 验证分析初步鉴定了猪腹股沟/阴囊疝的易感位点和 2 个位置功能候选基因: *UST* 和 *CDKAL1*, 该结果为进一步解析腹股沟/阴囊疝的分子遗传基础提供了新的思路。

**关键词:** 猪; 腹股沟/阴囊疝; 全基因组关联分析; 易感基因座位

中图分类号: S828; S813.3

文献标志码: A

文章编号: 0366-6964(2014)11-1775-09

## Genome-wide Association Study on Identifying Susceptibility Loci for Pig Inguinal/scrotal Hernia

RUAN Guo-rong<sup>2</sup>, LONG Yi<sup>1</sup>, SU Ying<sup>1</sup>, ZHANG Zhi-yan<sup>1</sup>, XIAO Shi-jun<sup>1</sup>,  
LIAO Xin-jun<sup>1,3</sup>, AI Hua-shui<sup>1</sup>, YANG Bin<sup>1</sup>, DENG Zheng<sup>1</sup>, XIN Wen-shui<sup>1</sup>,  
TANG Jian-hong<sup>1</sup>, REN Jun<sup>1</sup>, DING Neng-shui<sup>1\*</sup>, HUANG Lu-sheng<sup>1</sup>

(1. Candidate of National Key Laboratory for Animal Biotechnology, Jiangxi Agricultural University, Nanchang 330045, China; 2. Fujian Vocational College of Agriculture, Fuzhou 350119, China;  
3. College of Life Science of Jinggangshan University, Ji'an 343009, China)

**Abstract:** To identify susceptibility genes susceptible for pig inguinal/scrotal hernia and further reveal the pathogenesis mechanisms of the disorder, 734 commercial purebred pigs (237 cases in 220 trio families) containing Landrace and Large White breeds were collected and genotyped with Illumina Porcine 60k SNP chips. The genome-wide association studies (GWAS) based on the single-marker case-control model was used to identify SNPs related to pig inguinal/scrotal hernia.

收稿日期: 2014-03-26

基金项目: 国家自然科学基金(31071089); 教育部新世纪优秀人才计划资助项目(教育函[2012]80); 福建省教育厅科技计划项目(JB12380)

作者简介: 阮国荣(1965-), 男, 江西上犹人, 博士, 教授, 主要从事动物遗传育种与繁殖的研究, E-mail: rgr001@126.com; 龙毅(1987-), 男, 湖南湘乡人, 博士, 主要从事动物遗传育种与繁殖的研究, E-mail: longyilongyily@163.com。二人并列第一作者

\* 通信作者: 丁能水, 博士, 研究员、博士生导师, 主要从事动物育种新技术研究, E-mail: dingyd2005@hotmail.com

Four significant SNPs in a region of 19.31-19.59 Mb on SSC1 and a SNP at 17.05 Mb on SSC7 which showed significant association with pig inguinal/scrotal hernia. The *UST* gene within the susceptible region on SSC1 was responsible for biosynthesis of dermatan sulfate glycosaminoglycan that had been shown to be associated with human mucopolysaccharidosis accompanying several type of hernia. *CDKAL1* corresponding to the SSC7 locus might affect the occurrence of hernia through the function on cell death. Only MARC0077275 in *CDKAL1* was verified using transmission disequilibrium test (TDT) in an expanded population including 1 134 individuals containing 367 individuals with inguinal/scrotal hernia. In summary, this study identified 5 susceptible loci and 2 candidate genes for pig inguinal/scrotal hernia by GWAS, one of which was also verified in a larger population by TDT analysis. These findings provide novel insight into the molecular and genetic basis of inguinal/scrotal hernia in pigs.

**Key words:** swine; inguinal/scrotal hernia; genome-wide association study; susceptibility loci

腹股沟/阴囊疝是猪生产中最常见的生长发育缺陷之一,主要表现为腹内容物由单侧或双侧腹股沟裂口直接脱至腹股沟外侧的皮下或阴囊内,造成局部膨胀突起。其发病率一般为 0.11%~10%<sup>[1-5]</sup>,给世界养猪业造成了较大经济损失且影响猪只福利<sup>[5]</sup>。该疾患不仅妨碍猪只生长发育,损害种猪使用性能(一般不留作种用),降低商品猪价值,而且可能因发生疝囊粘连和嵌顿疝而增加猪只死亡率。猪腹股沟/阴囊疝的发生是遗传和环境因素共同作用的结果,其家族聚集性发病和遗传率在不同品种间及不同遗传背景下的差异性提示遗传因素对其发病起到重要作用。在育种实践中,通常将患病个体及其同胞甚至父母本全部淘汰,不留作种用。但这种单纯的表型选择难以降低疝气的发生率,且可能丧失优秀的种质资源。因此,分离和鉴别猪腹股沟/阴囊疝的易感位点及其主要致病基因成为猪腹股沟/阴囊疝抗病育种的关键。

迄今为止, pigQTLdb (<http://www.animalgenome.org/cgi-bin/QTLdb/SS>) 中报道的猪腹股沟/阴囊疝基因组易感 QTLs 共有 45 个<sup>[6-7]</sup>。其中, E. Grindflek 等<sup>[6]</sup> 利用 137 个微卫星标记对包含 103 个疝气同胞对的 282 个长白猪样品进行基因组扫描,将影响猪腹股沟/阴囊疝的易感区域定位到 SSC1、SSC2、SSC5、SSC6、SSC7、SSC15、SSC17 和 X 染色体上;本实验室 N. Ding 等<sup>[7]</sup> 在白色杜洛克×二花脸 F2 资源家系中利用 194 个微卫星标记在 SSC2、SSC4、SSC8、SSC10、SSC13、SSC16 和 SSC18 上发现猪腹股沟/阴囊疝易感 QTLs。F. X. Du 等<sup>[8]</sup> 在商业猪种 SSC2 上 3.42 和 65 cM 以及 SSC12 上 20 和 67 cM 处共鉴别到 5 个阴囊疝候选易感区域。Z. Q. Du 等<sup>[9]</sup>

的深入研究进一步证实了这一结果。但这些腹股沟/阴囊疝易感区域及候选基因的分子作用机制仍不清楚,综合人类和模式生物疝气解剖学、病理学、胚胎学和遗传及分子生物学的研究成果,推测腹股沟/阴囊疝的发生可能与鞘状突闭合失败(细胞增殖和凋亡异常)<sup>[10-13]</sup> 及胶原代谢紊乱<sup>[14-17]</sup> 相关。

GWAS 自产生以来在人类复杂疾病研究中得到迅速发展,至今人类全基因组关联分析数据库 ([http://www.genome.gov/GWA\\_Studies/](http://www.genome.gov/GWA_Studies/)) 中公开发表的 GWAS 研究成果共 1 844 项,其中报道了 12 802 个遗传变异与 700 余种疾病或性状显著关联,可见 GWAS 在鉴别疾病易感位点或候选基因的研究中成为强有力的工具。但有关猪腹股沟/阴囊疝的 GWAS 研究尚未有正式报道,本研究拟在商业纯种长白和大白猪腹股沟/阴囊疝三联体核心家系群体中开展全基因组关联分析并利用大样本家系群进行 TDT 验证,从而鉴别猪腹股沟/阴囊疝易感位点及候选基因,为进一步精细定位致病基因或因果突变、建立剔除猪腹股沟/阴囊疝遗传缺陷的分子育种技术提供了科学依据。

## 1 材料与方法

### 1.1 病例诊断及样品收集方法

病例诊断:腹股沟/阴囊疝的诊断均由专业兽医或育种人员严格按照疝气临床诊断标准<sup>[18]</sup> 进行。对公猪触诊,阴囊疝个体常表现为腹内容物沉降至阴囊,引起一侧或双侧的阴囊增大,患侧阴囊硬度不一,可触摸到阴囊疝内容物,如提举两后肢,常可使疝内容物回至腹腔而使阴囊缩小,但放下后或腹压加大后又恢复原发病症状。对母猪可确诊为腹股沟疝。

样品收集:样品采集包括腹股沟/阴囊疝患病个体及其父母本或同性别同胞(三联体核心家系—Trios),其中患病个体均采集照片,所有样品按要求详细登记品种、耳号、性别、系谱及患病状态等信息。

## 1.2 样本的来源和数量

GWAS 群体:本研究从全国 8 个省(市)16 个大

型原种猪场采集的三联体核心家系包含 237 个患病个体在内的 734 个样本(表 1),其中长白 240 头(患病个体 81 头),大白 494 头(患病个体 156 头)。样品中包含 2 个个体以上的家系 220 组,3 个个体以上的家系 176 组。每个半同胞家系个体总数从 2 到 23 个不等,患病个体 1~16 个不等。

表 1 长白和大白猪群体中腹股沟/阴囊疝患病个体和非患病个体数量分布情况

Table 1 The number of affected and unaffected pigs in Landrace and Large White pigs

品种 Breed	总数/头 Total number	患病总数/头 Affected number	父母及同胞总数/头 Parents and sibs number
长白 Landrace	240	81	159
大白 Large White	494	156	338
合计 Total number	734	237	497

验证群体:为了验证 GWAS 所得到的结果,本研究又新增了 400 个个体(130 个患病个体),并与上文中 734 个个体合并进行 TDT 验证分析。验证群体中长白共有 366 个个体(包含 123 个腹股沟/阴囊疝患病个体),大白共有 768 个个体(包含 244 个腹股沟/阴囊疝患病个体)。

## 1.3 基因组 DNA 提取

采集猪耳组织 0.5~1.0 g,按照标准酚/氯仿抽提、乙醇沉淀法提取 DNA,将其溶于 TE 缓冲液中,置于-20 °C 冰箱保存备用。用 Nanodrop 1000 全波长紫外/可见光扫描分光光度计(Thermo Scientific, USA)测定 DNA 的浓度和质量: $A_{260\text{ nm}}/A_{280\text{ nm}}$  的比值为 1.8~2.0,  $A_{260\text{ nm}}/A_{230\text{ nm}}$  比值为 1.5~2.3 视为合格。将浓度统一稀释至  $50\text{ ng} \cdot \mu\text{L}^{-1}$ ,用 0.8% 琼脂糖凝胶电泳检测 DNA 片段的完整性,片段长度大于 10K,无 RNA 和蛋白污染则判定为合格。质量合格的 734 个 DNA 样品用于 Illumina 猪 60K SNP 芯片进行基因分型。400 个新增样品 DNA 利用 Taqman 探针法进行 SNP 基因分型。

## 1.4 基因分型

Illumina 猪 60K SNP 芯片基因分型:将质量合格的 734 个纯种商业猪种腹股沟/阴囊疝群体的 DNA 样品送至北京怡美通德科技发展有限公司,利用 Illumina Bead Array SNP 分型平台进行 Illumina 猪 60K SNP 芯片(San Diego, CA, USA)扫描,并利用 BEAD-STUDIO version 3.2(Illumina, USA)软件进行基因分型。

Taqman 探针 SNP 基因分型:根据 Ensemble

公布的猪基因组序列(Sscrofa10.2)中 5 个猪腹股沟/阴囊疝候选 SNP 位点所在区域的基因组序列,通过 Primer Express V 3.0 软件来设计 5 对 Taqman 探针和 5 对与之对应的引物(表 2),交由上海英骏生物技术有限公司合成。探针 1 的 5' 端荧光标记 FAM,探针 2 的 5' 端荧光标记 VIC,3' 端荧光标记均为 MGB。Taqman 探针 SNP 基因分型扩增体系均由美国应用生物系统公司(ABI, USA)提供,10  $\mu\text{L}$  的反应体系中含 DNA 模版 20 ng(1  $\mu\text{L}$ ),超纯水 3.4  $\mu\text{L}$ ,2×Taqman Genotyping Master Mix 5  $\mu\text{L}$ ,10 mmol·L<sup>-1</sup> 上下游引物各 0.2  $\mu\text{L}$ ,10 mmol·L<sup>-1</sup> 探针 1 和探针 2 各 0.1  $\mu\text{L}$ 。利用 7900HT fast 实时定量 PCR 仪(ABI, USA)进行 PCR 反应,反应条件为 50 °C 2 min;95 °C 10 min;(95 °C 15 s,各引物退火及延伸温度 60 °C 1 min)×40 个循环。PCR 反应和荧光信号读取均在 ABI7900HT Sequence Detector(384 孔板)上完成,以 SDS2.4 图像分析软件 Allelic Discrimination 程序进行终点分析,除去本底信号,通过检测不同等位基因所标的 FAM 和 VIC 荧光强度,判断各待测样本基因型为野生型纯合子、突变型纯合子还是杂合子,分析后自动分簇给出判型结果。

## 1.5 GWAS 数据质量控制

为验证样本采集及系谱记录的准确性,本研究利用 plink(v1.07)软件对所有个体进行性别再检验, $F$  值小于 0.3 则为雌性个体, $F$  值大于 0.7 则为雄性个体。 $F$  值介于 0.3~0.7,则通过 X 染色体上特异区域的基因型比较来判别。根据孟德尔分离规

表 2 用于 Taqman 探针 SNP 基因分型的探针和引物

Table 2 Probes and primers used for Taqman SNP genotyping

标记 Marker	引物序列(5'-3') Primer sequence	探针序列(5'-3') Probe sequence	退火温度/℃ Annealing temperature	扩增片段 大小/bp Size
ASGA0001418	FP:CACTAGCAGCTCTGCCTTGA RP:GTCCATTGATATGTTTCCACA	P1:TTAACCAGCCCCCTGAGA P2:TTAACCAGTCCCCTGA	60	159
ASGA0001422	FP:CCTGGATGTTCTCTCTGGAATGT RP:CTTGCATGAACAGCTGCTTCA	P1:AATCCTTGAAGGGCT P2:CCTTGAAGGACTTTG	60	86
H3GA0001012	FP:AACCTGCCTTCTTGAACGCC RP:GCCAGCAAATACAAAACCC	P1:TTCAGGCAGACTCA P2:TTCAGGCGGACTCA	60	130
ASGA0097745	FP:AAGCCCAGAAACAGGCAGTG RP:AGCCTGGAACATAAACATAA	P1:CCTTACGATTCAAGGTA P2:ACCTTACGATTCAAGATA	60	169
MARC0077275	FP:TTAAGGAGATCATCTGACGCAGTG RP:CTGGGCTAGAAGCCAGGGAT	P1:AAGTGGCAGAATCA P2:AAGTGGCAGAACCA	60	21

律,自主编写 perl 脚本判别亲子关系,然后根据共有父本和母本进行家系合并。

利用 plink 软件对所有样本的 60K SNP 芯片扫描分型数据进行质量控制,标准:剔除 SNP 检出率小于 90% 的个体,剔除 SNP 位点检出率小于 90%、哈迪-温伯格平衡检验  $P \leq 0.001$  (卡方检验)、次等位基因的频率  $MAF \leq 0.05$  和性染色体上及无法定位的 SNP 标记。

## 1.6 统计分析

1.6.1 病例-对照全基因组关联分析 本研究采用基于 SNP 的病例-对照全基因组关联分析,通过比较全基因组上每一个标记位点上患病个体和健康个体中等位基因及基因型频率的差异,搜寻与控制目标性状的基因座位处于连锁不平衡的位点。为了校正由于亲缘关系以及群体分层效应对关联分析造成可能的假阳性影响,本研究采用广义线性混合模型加以控制,通过 GenABEL 软件包<sup>[19]</sup>进行全基因组关联分析。广义线性混合模型:

$$\eta = X\beta + S\alpha + Qv + Zu + e$$

$\eta = \ln(P/1-P)$ ,  $\eta$  为符合伯努利分布的表型向量,  $P$  为群体中每个个体患阴囊疝的平均可能性。  $X$  为固定效应的指示矩阵,  $\beta$  为固定效应的估计参数;  $S$  为 SNP 的指示矩阵,  $\alpha$  为 SNP 的效应;  $Q$  为个体占每个亚群的比例,  $v$  为群体结构的效应;  $Z$  为随机效应的指示矩阵;  $u$  为预测的随机效应,随机效应的方差分布为  $2K * V_g$ , 其中  $K$  为校正后亲缘关系

矩阵(由全基因组标记基因型计算而得),  $V_g$  为利用限制最大似然法(REML)估算的方差;  $e$  为残差。

Q-Q plot: 利用 R 软件中自主编写的分位数图(Quantile-quantile plot, Q-Q plot)对群体分层效应进行评估,通过比较观测  $-\log_{10} P$  与零假设(假设无 SNP 位点是腹股沟/阴囊疝的易感位点)条件下的预期  $-\log_{10} P$  的分布曲线来估测是否存在群体分层效应及真正关联的 SNP 位点。若观测曲线趋于预期置信度,则群体分层效应得到很好控制,若观测曲线线尾显著高于预期曲线度,则该 SNP 位点可能为腹股沟/阴囊疝易感位点。

1.6.2 传递不平衡检验(TDT) 基于家系的传递不平衡检验(TDT)是研究疾病性状的重要方法,可有效避免复杂的群体分层效应导致的结果偏差。本研究采样来源广泛,遗传背景复杂,具有一定程度的群体分层效应,且新增个体没有 60K SNP 数据不能通过全基因组数据进行群体分层效应的校正,因而选用 plink v 1.07 程序中的 TDT<sup>[20]</sup> 命令进行关联分析。其原理:分析某个等位基因从杂合子父母传递给患病子代的机率是否高于预期值(50%),如果这个家系至少有一个亲本是杂合子,则进行经典的 TDT 分析;如果双亲的基因型缺失,则进行同胞 TDT(Sib transmission disequilibrium test, S-TDT) 检验<sup>[21-22]</sup>。当连锁不平衡程度很高的情况下,标记与致病基因紧密连锁时,它的检测效率很高。其计算公式:

$$TDT_{Max} =$$

$$Max \left( \frac{(b_1 - c_1)^2}{b_1 + c_1}, \frac{(b_2 - c_2)^2}{b_2 + c_2}, \dots, \frac{(b_n - c_n)^2}{b_n + c_n} \right)$$

其中: $n$  是 SNP 的数目, $b_i$  为传递的 SNP 或单倍型数目, $c_i$  为未传递的 SNP 或单倍型的数目。

## 2 结果

### 2.1 GWAS 统计分析方法基因分型数据的质控结果

734 个样本的 Illumina 猪 60K SNP 芯片(包含 62 163 个 SNPs 位点)基因分型结果经质量控制后,剩余 708 个个体的 36 574 个 SNPs 标记可用于后续与猪腹股沟/阴囊疝的全基因组关联分析。

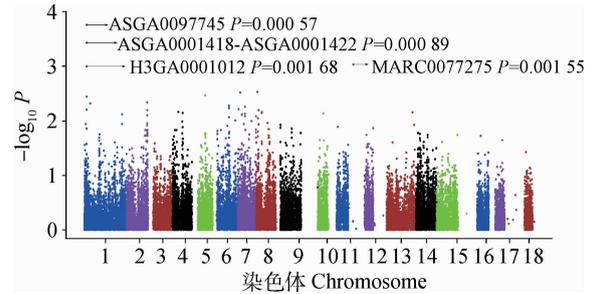
### 2.2 病例-对照全基因组关联分析结果

通过 GenABEL 软件对所有样品进行病例-对照全基因组关联分析(图 1),并经 Q-Q plot 检验(图 2),结果 Q-Q 图尾部有 5 个 SNPs 位点显著偏离零假设分布,表明存在与猪腹股沟/阴囊疝显著关联位点(图 1)。由图 2 可见,当使用混合模型中的 Q+k 模型分析时,存在过度校正的问题,当根据观测值的斜率重新校正后,可以看出有 5 个标记(表 3)超出了期望值,达到基因组显著水平。其中,在 SSC1 和 SSC7 上分别有 4 个和 1 个易感位点,SSC1 上的 SNP ASGA0097745 与猪腹股沟/阴囊疝关联最为显著( $P=0.00057$ )。

### 2.3 验证群体 TDT 分析结果

在扩大群体至 1 134 个个体(包含 367 个腹股沟/阴囊疝患病个体)的腹股沟/阴囊疝三联体家系群中进行 TDT 分析,结果显示:5 个 GWAS 揭示的易感位点中仅 SSC7 上 *CDKAL1* 基因内的 MARC0077275 标记与腹股沟/阴囊疝的关联性达

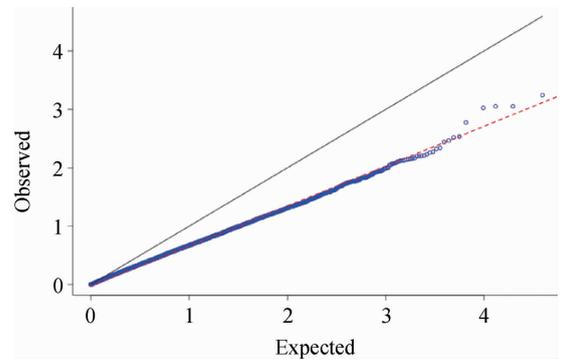
到极显著水平( $P<0.01$ )(表 4),该位点 TT 基因型的个体更易患腹股沟/阴囊疝(表 5),突变型 T 基因为不利突变。



横坐标为染色体,纵坐标为关联分析  $P$  值对数的相反数  
Abscissa axis means the chromosome,ordinate axis means the opposite of exponential  $P$  value

图 1 曼哈顿图(Q+K 模型,GenABEL 分析)

Fig.1 Manhattan plot(Q+K model,GenABEL program)



横坐标为预期  $P$  值对数的相反数,纵坐标为计算  $P$  值对数的相反数;虚线为校正后计算值的置信线

Abscissa axis means the expected value of the opposite of exponential  $P$  values,ordinate axis means the calculated value of the opposite of exponential  $P$  values;The dashed line indicates corrected expected confidence line

图 2 Q-Q 图

Fig.2 Q-Q plot

表 3 病例-对照全基因组关联分析鉴别的猪腹股沟/阴囊疝易感 SNPs 标记位点

Table 3 Significant SNPs identified by SNP-based case-control GWAS on pig inguinal/scrotal hernia

染色体 Chromosome	标记号 SNP ID	物理位置/bp Position	最近基因 Nearest gene	$P$ 值 $P$ value
1	ASGA0097745(C/T)	19585039	121.85 <i>DUST</i>	0.00057
1	ASGA0001418(A/G)	19345101	<i>UST</i>	0.00089
1	ASGA0001422(T/C)	19312605	<i>UST</i>	0.00089
1	H3GA0001012(C/T)	19460835	<i>UST</i>	0.00168
7	MARC0077275(C/T)	17049497	<i>CDKAL1</i>	0.00155

D. 代表这个标记在基因的下游(downstream)

D. Means downstream of the gene

表 4 关联 SNP 位点在验证群体中单点 TDT 重复验证结果

Table 4 Verification result of TDT based on single SNP marker in verification populations

染色体 Chromosome	多态位点 SNP	位置/bp Location	等位基因 Allele	等位基因频率 Allele frequency	传递 T	未传递 U	卡方 Chisq	P 值 P value
7	MARC0077275(C>T)	17049497	C	0.421	60	93	7.118	0.007 6**
			T	0.579	93	60		
1	ASGA0097745(A>G)	19585039	A	0.827	30	34	0.25	0.617 1
			G	0.173	34	30		
1	ASGA0001422(T>C)	19312605	T	0.826	36	40	0.210 5	0.646 4
			C	0.174	40	36		
1	ASGA0001418(A>G)	19345101	A	0.823	37	40	0.116 9	0.732 4
			G	0.177	40	37		
1	H3GA0001012(C>T)	19460835	C	0.806	42	42	0	1
			T	0.194	42	42		

\*\* . 表示 TDT 分析中与腹股沟/阴囊疝的关联性达极显著水平 ( $P < 0.01$ )

\*\* . Means extreme significance in TDT analysis ( $P < 0.01$ )

表 5 MARC0077275 位点各基因型在不同品种分布情况

Table 5 The genotype distribution of MARC0077275 in Landrace and Large White pigs

品种 Breed	基因型 Genotype	数目/头 Number	患病数目/头 Case number	比例/% Ratio
大白 Large White	TT	371	131	35
	TC	302	92	30
	CC	65	13	20
长白 Landrace	TT	75	28	37
	TC	147	48	33
	CC	144	46	32

### 3 讨论

作为一种复杂的遗传缺陷疾病,人们对腹股沟/阴囊疝的认识经历了从不完全显性单基因<sup>[23]</sup>到双隐性基因<sup>[24]</sup>,再到复杂多基因背景下的主效基因的一系列过程。迄今为止,所有对猪腹股沟/阴囊疝易感位点定位的公开研究结果均是基于低密度标记扫描而获得,本研究利用高密度 Illumina 猪 60K SNP 芯片数据,对商业纯种长白和大白腹股沟/阴囊疝核心家系群体进行 GWAS 分析鉴别到疝气易感位点,为猪腹股沟/阴囊疝抗病育种提供了科学依据和理论借鉴。

#### 3.1 病例-对照全基因组关联分析

本研究利用商业纯种长白和大白腹股沟/阴囊疝核心家系群体共 734 个个体(含 237 个患病个体)的 Illumina 猪 60K SNP 芯片分型数据,通过基于病例-对照的全基因组关联分析分别在 SSC1 上的 19.31~19.59 Mb 区域和 SSC7 上 17.05 Mb 处鉴别到显著影响猪腹股沟/阴囊疝发生的易感位点/区域。通过比较发现,E. Grindflek 等<sup>[6]</sup>在 SSC1 上定位到的 2 个显著影响猪腹股沟/阴囊疝发生的易感位点 SW1515~SW1332 位于 14.88~21.06 Mb 区域,与本研究在 SSC1 上所鉴别的易感区域(19.31~19.59 Mb)相互间有重叠,此区域可作为

研究猪腹股沟/阴囊疝的重要候选区域。而本研究在 SSC7 的易感位点/区域与 E. Grindflek 等<sup>[6]</sup>在同一染色体上定位到 2 个显著影响猪腹股沟/阴囊疝发生的易感位点/区域(分别为 27.56 Mb 的 TNFB 标记处和 12.38 Mb 的 S0212 标记处)均不一致,因而本研究在 SSC7 上鉴别到 1 个新的猪腹股沟/阴囊疝易感位点(MARC0077275),推测猪腹股沟/阴囊疝的发病是由复杂多基因控制的,需要进一步扩大样本继续研究。

### 3.2 验证群体中 TDT 分析

对病例-对照 GWAS 分析发现的 5 个易感 SNPs 位点在更大规模的长白和大白腹股沟/阴囊疝三联体家系群体中进行更为严格的 TDT 重复验证分析。结果显示:5 个 SNPs 位点中仅位于 *CDKAL1* 内的 MARC0077275 位点在扩大群体中与腹股沟/阴囊疝的发生呈极显著相关( $P < 0.01$ )。该结果进一步证实病例-对照 GWAS 揭示的易感位点,但有些 SNP 位点没有得到验证,其原因可能是统计方法不同:采用病例-对照 GWAS 分析在样本量适当的情况下更能够揭示疾病易感位点,TDT 分析以家系为基础,因而能避免群体结构分层效应。但其要求亲代的标记位点等位基因型是杂合型,以观察可能的易感标记等位基因传递给患病子代的概率,对数据的要求更为严格。对于疾病性状,一般采用病例-对照研究揭示群体水平上某标记位点与疾病间存在某种关联性的前提下,再进行 TDT 分析,进一步证实病例对照的结果<sup>[25]</sup>。本研究通过大样本量的 TDT 验证证实了 *CDKAL1* 基因中 MARC0077275 标记与腹股沟/阴囊疝呈极显著关联,因而 *CDKAL1* 基因可作为腹股沟/阴囊疝易感的重要候选基因。

### 3.3 腹股沟/阴囊疝位置候选易感基因

根据 GWAS 分析及位置候选基因搜寻结果,在 SSC1 上易感区域(19.31~19.59 Mb)内仅有 1 个已注释的基因:糖醛基-2-磺基转移酶(Uronyl-2-sulfotransferase,*UST*)。而 SSC7 上易感位点则位于或邻近细胞周期蛋白依赖激酶样 1(Cyclin-dependent kinase-like 1,*CDKAL1*)基因。本研究拟定该 2 个基因为腹股沟/阴囊疝的位置功能易感基因。

*UST* 的主要作用是催化转移硫酸基至糖醛基残基 2 号位置上,是一类重要的粘多糖(Glycosamino glycans, GAG)合成调控酶。硫酸皮肤素

(Dermatan sulfate, DS)和硫酸软骨素(Chondroitin sulfate, CS)均属于重要的粘多糖,广泛存在于体内各种组织中,是细胞外基质(ECM)和细胞表面的重要组成成分。研究表明,DS/CS 作为蛋白粘多糖的核心功能单元,可以通过硫酸氨基葡聚糖(Glycosamino glycans, GAGs)链和多种生长因子、细胞趋化因子及胞外基质分子等相结合作用于细胞的生长和增殖,还可以受体/配体结合形式通过糖基化磷脂酰肌醇和钙离子的第二信号系统调节细胞内激酶磷酸化过程,影响细胞的蛋白合成和生理功能<sup>[26]</sup>。此外,CSPGs 还可通过细胞内通路传导细胞内信号,可表达细胞凋亡相关分子,引发此类分子的低聚反应,诱发细胞凋亡<sup>[27]</sup>。硫酸皮肤素和硫酸软骨素等粘多糖的代谢异常,造成在体内的过多聚集可引发多种粘多糖症(Mucopolysaccharidosis, MPS)<sup>[28]</sup>。根据临床表现和酶缺陷,MPS 可以分为 I~VII 等。除 II 型汉特综合症(Hunter syndrome)为性连锁隐性遗传外,其余均属常染色体隐性遗传病。这些综合症经常可见各类疝气表型症候的出现。鉴于本研究鉴别出的 SSC1 上猪腹股沟/阴囊疝易感 SNP 位点位于 *UST* 基因内,推测 *UST* 基因可能通过调控硫酸皮肤素、硫酸软骨素等粘多糖的功能模块或代谢通路,导致这些粘多糖的功能或代谢异常,从而引发疝气。所以,*UST* 作为猪腹股沟/阴囊疝的位置候选基因值得进一步深入研究。

而 SSC7 上的易感位点在验证群体 TDT 分析中也达到显著关联水平,其位于细胞周期蛋白 5 调节亚基相关蛋白 1 类似物 1(*CDK5* regulatory subunit associated protein 1-like 1,*CDKAL1*)基因内,编码一种甲硫基转移酶。尽管 *CDKAL1* 编码蛋白的功能目前尚不清楚,但推测其与细胞周期素依赖激酶 5 调控亚基相关蛋白 1(*CDK5RAP1*)的功能类似,即通过结合到 *CDK5* 的活化因子 p35 上抑制 *CDK5* 的活性<sup>[29]</sup>。所以,若 *CDKAL1* 基因发生变异,则对 *CDK5* 失去抑制作用。人类 GWAS 研究发现 *CDKAL1* 与 2 型糖尿病有关<sup>[30]</sup>。此外,小鼠 *CDKAL1* 基因敲除试验证明,*CDKAL1* 基因可影响  $Ca^{2+}$  活性通路<sup>[31]</sup>,而疝气发生的分子病理机制中鞘状突闭合失败与钙离子的释放及储存有关,因此推测其可能通过调控钙离子通路而作用于细胞凋亡过程,最终影响到鞘状突闭合,引发疝气发生。但 *CDKAL1* 基因敲除的小鼠是否会导致疝气表型并未有报道,其可能影响阴囊/腹股沟疝的分子机制有

待深入研究。

#### 4 结 论

本研究利用猪 60K SNP 基因型数据,采用病例-对照全基因组关联分析鉴别到 5 个与猪腹股沟/阴囊疝存在显著相关的易感位点/区域及 2 个位置功能候选基因。为腹股沟/阴囊疝的遗传研究奠定了分子学基础。

致谢:本试验腹股沟/阴囊疝样本的采集得到江西正邦集团辖属种猪场、广西农垦永新畜牧有限公司良圻原种猪场、广西杨翔农牧有限公司、广东华农温氏畜牧股份有限公司、河南正阳诸美原种猪场、海南罗牛山种猪育种有限公司、安徽长风农牧科技有限公司、福建省永诚华多种猪有限公司、山东省日照原种猪场等全国 16 个原种猪场的大力支持,在此一并致谢!

#### 参考文献:

[1] MATSSON P. Prevalence of congenital defects in Swedish Hampshire, Landrace and Yorkshire pig breeds and opinions on their prevalence in Swedish commercial herds[D]. Swedish:Swedish University of Agricultural Sciences,2011.

[2] PARTLOW G D, FISHER K, PAGE P D, et al. Prevalence and types of birth defects in Ontario swine determined by mail survey[J]. *Can J Vet Res*, 1993, 57(2):67.

[3] 陈其云. 小公猪阴囊疝的诊治[J]. 畜牧兽医科技信息, 2008, 4:79.

[4] 范国英, 刘俊伟, 钟 华, 等. 某规模化猪场公猪腹股沟阴囊疝发病及治疗情况[J]. 养猪, 2011, 2:87.

[5] WALTERS J R. Have we forgotten about inherited disease? [C]. In AGBU Pig Genetics Workshop, 2010:79186.

[6] GRINDFLEK E, MOE M, TAUBERT H, et al. Genome-wide linkage analysis of inguinal hernia in pigs using affected sib pairs[J]. *BMC Genet*, 2006, 7(1): 25.

[7] DING N, MAO H, GUO Y, et al. A genome-wide scan reveals candidate susceptibility loci for pig hernias in an intercross between White Duroc and Erhualian[J]. *J Anim Sci*, 2009, 87(8):2469-2474.

[8] DU F X, MATHIALAGAN N, DYER C, et al. Mapping genes affecting scrotal hernia condition in domes-

tic pigs[J]. *J Dairy Sci*, 2004, 87:453.

- [9] DU Z Q, ZHAO X, VUKASINOVIC N, et al. Association and haplotype analyses of positional candidate genes in five genomic regions linked to scrotal hernia in commercial pig lines[J]. *PLoS ONE*, 2009, 4(3): e4837.
- [10] TANYEL F C, ERDEM S, BÜYÜKPAMUKÇU N, et al. Smooth muscle within incomplete obliterations of processus vaginalis lacks apoptotic nuclei[J]. *Urol Int*, 2002, 69(1):42-45.
- [11] MOURAVAS V, KOLETSA T, SFOUGARIS D, et al. Smooth muscle cell differentiation in the processus vaginalis of children with hernia or hydrocele[J]. *Hernia*, 2010, 14(2):187-191.
- [12] TANYEL F C, ULUSU N N, TEZCAN E F, et al. Total calcium content of sacs associated with inguinal hernia, hydrocele or undescended testis reflects differences dictated by programmed cell death[J]. *Urol Int*, 2003, 70(3):211-215.
- [13] BEUERMANN C, BECK J, SCHMELZ U, et al. Tissue calcium content in piglets with inguinal or scrotal hernias or cryptorchidism[J]. *J Comp Pathol*, 2009, 140(2):182-186.
- [14] BENDAVID R. The unified theory of hernia formation[J]. *Hernia*, 2004, 8(3):171-176.
- [15] RODRIGUES C J, YOO J H, RODRIGUES JUNIOR A J. Elastin (ELN) gene point mutation in patients with inguinal hernia[J]. *Genet Mol Biol*, 2006, 29(1):45-46.
- [16] AJABNOOR M, MOKHTAR A, RAFEE A, et al. Defective collagen metabolism in Saudi patients with hernia[J]. *Ann Clin Biochem*, 1992, 29(4):430-436.
- [17] PANS A. New prospects in the etiology of groin hernias[J]. *Chirurgie; memoires de l'Academie de chirurgie*, 1999, 124(3):288-297.
- [18] 寇向龙. 猪腹股沟阴囊疝治疗初探[J]. 甘肃科技, 2012, 28(4):139-142.
- [19] AULCHENKO Y S, RIPKE S, ISAACS A, et al. An R library for genome-wide association analysis[J]. *Bioinformatics*, 2007, 23(10):1294-1296.
- [20] 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.
- [21] WHITTEMORE A S, HALPERN J. A class of tests for linkage using affected pedigree members[J]. *Biometrics*, 1994, 50(1):118-127.

- [22] RISC N, MERIKANGAS K. The future of genetic studies of complex human diseases[J]. *Science*, 1996, 273(5281):1516-1517.
- [23] BERGE S. The inheritance of paralysed hind legs, scrotal hernia and atresia ani in pigs[J]. *J Hered*, 1941, 32(8):271-274.
- [24] HUSTON R, SAPERSTEIN B, SCHONEWEIS D, et al. Congenital defects in pigs[J]. *Vet Bulletin (UK)*, 1978, 48:645-675.
- [25] BENNETT S, LUCASSEN A, GOUGH S, et al. Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus[J]. *Nat Genet*, 1995, 9(3):284-292.
- [26] MAEDA N, FUKAZAWA N, HATA T. The binding of chondroitin sulfate to pleiotrophin/heparin-binding growth associated molecule is regulated by chain length and oversulfated structures[J]. *J Biol Chem*, 2006, 281(8):4894-4902.
- [27] FUJIMOTO T, KAWASHIMA H, TANAKA T, et al. CD44 binds a chondroitin sulfate proteoglycan, aggrecan[J]. *Int Immunol*, 2001, 13(3):359-366.
- [28] LEHMAN T J, MILLER N, NORQUIST B, et al. Diagnosis of the mucopolysaccharidoses[J]. *Rheumatology*, 2011, 50(suppl 5):v41-v48.
- [29] CHING Y P, PANG A S, LAM W H, et al. Identification of a neuronal Cdk5 activator-binding protein as Cdk5 inhibitor [J]. *J Biol Chem*, 2002, 277(18):15237-15240.
- [30] STEINTHORSDOTTIR V, THORLEIFSSON G, REYNISDOTTIR I, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes [J]. *Nat Genet*, 2007, 39(6):770-775.
- [31] OHARA-IMAIZUMI M, YOSHIDA M, AOYAGI K, et al. Deletion of CDKAL1 affects mitochondrial ATP generation and first-phase insulin exocytosis [J]. *PLoS ONE*, 2010, 5(12):e15553.

(编辑 郭云雁)