

应用 MGB 荧光探针技术确定猪 *KIT* 基因拷贝重复数

王 韬, 唐 辉*, 曾勇庆, 李晓宁, 闫 超, 董林松, 于希江*

(山东农业大学动物科技学院, 泰安 271018)

摘要: 旨在建立一种简便、快速检测猪 *KIT* 基因拷贝重复数(CNVs)的方法并应用于未知样品的检测。在 *KIT* 外显子 2 中设计 TaqMan-MGB 探针和引物, 建立荧光实时定量 PCR 检测 *KIT* 基因 CNVs 的方法。结果表明, TaqMan-MGB 探针扩增 *KIT* 基因目的片段, 不同 DNA 梯度浓度 Ct 值差异极显著 ($P < 0.001$), 变异系数低 (0.12%~0.26%); *KIT* 与内标基因 *ESR* 的扩增效率相差只有 2.4%; 采用聚类分析推测了待测的 50 枚样品 *KIT* 的 CNVs 分别归属于 2、3、4、5、6 个拷贝。本研究建立的 TaqMan-MGB 探针实时荧光定量 PCR 检测 *KIT* 基因 CNVs 的方法, 具有简便、快速、特异性和稳定性好的特点, 适合于普通实验室应用。

关键词: 猪; 显性/*KIT* 位点; 拷贝数; MGB 探针; 荧光定量 PCR

中图分类号: S852.4

文献标识码: A

文章编号: 0366-6964(2011)03-0437-05

Detecting Copy Number Variations in Porcine *KIT* by TaqMan-MGB Probe

WANG Tao, TANG Hui*, ZENG Yong-qing, LI Xiao-ning, YAN Chao,

DONG Lin-song, YU Xi-jiang*

(College of Animal Science & Technology, Shandong Agricultural University,
Tai'an 271018, China)

Abstract: The main objective of the present study was to establish a simple, rapid method for detecting copy number variations (CNVs) in porcine *KIT* gene. TaqMan-MGB probes and primers were designed according to sequence of exon 2 of *KIT* gene, and a real-time fluorescence quantitative PCR procedure was established for the CNVs detection. The results showed that amplification curves of *KIT* obtained by TaqMan-MGB probe were a series of parallel curves corresponding to 2-fold serial dilutions of DNA samples, Ct values between groups were significantly different ($P < 0.001$), and the coefficient-of-variations were low (from 0.12% to 0.26%). The amplification efficiencies of the *KIT* and *ESR* were approximately equal. The CNVs in *KIT* of 50 pigs were estimated by cluster analysis, assigned to 2, 3, 4, 5 or 6 copies, respectively. The real-time quantitative PCR using TaqMan-MGB probe is a simple, rapid method with high resolution and stability to measure CNVs in *KIT*, and it could be carried out in common laboratories.

Key words: swine; dominant / *KIT* locus; copy number; MGB probe; FQ-PCR

猪的显性白被毛表型由 *KIT* 原癌基因控制^[1-2], 该基因编码干细胞生长因子受体 (the mast/stem cell growth factor receptor)^[3]。Marklund 等提出, 显性白表型由 *KIT* 基因的 2 个突变引起: 一

是基因拷贝数的串联重复 (~450 kb), 一是内含子 17 第一个碱基 G→A 的突变 (导致外显子 17 表达缺失)^[4]。

收稿日期: 2010-04-01

基金项目: 山东省中青年科学家奖励基金 (2008BS07009); 山东省科技攻关 (2009GG10009036); 国家 863 重大专项 (2008101002)

作者简介: 王 韬 (1984-), 男, 山东青岛人, 硕士生, 主要从事动物遗传育种研究, E-mail: wangtaosdau@gmail.com

* 通讯作者: 唐 辉, E-mail: tanghui@sda.u.edu.cn; 于希江, E-mail: xijiang.yu@gmail.com

从分子水平确定 *KIT* 的基因型有助于在育种实践中快速准确地控制猪的被毛类型,但第一步必须检测该基因拷贝数变异(copy number variations, CNVs)。在最近的研究中,实时定量 PCR^[5-6]、侵染检测(invader assay)^[7]、焦磷酸测序(pyrosequencing)^[7]、寡核苷酸连接分析法(qOLA)^[8]等技术都被用来进行 CNVs 检测。相对于其他方法复杂、不便于一般实验室操作的缺点,实时定量 PCR 通过荧光信号的检测可以直接对产物进行定量,操作简单、安全,自动化程度高,不易产生污染^[9]。在荧光定量检测中,TaqMan-MGB 探针比普通探针灵敏度更高,特异性更好。本试验采用 TaqMan-MGB 探针技术,通过荧光定量 PCR 方法测定猪 *KIT* 基因的拷贝重复数,有望能够建立一种更加有效的检测方法。

1 材料与与方法

1.1 试验材料

中系白色杜洛克猪 38 头,由临沂江泉原种猪场和山东农业大学合作培育;杜洛克猪 6 头,毛色棕红;大白猪 6 头,毛色纯白。所有猪只均来自山东临沂江泉原种猪场,采取耳组织样品用于 DNA 提取。

1.2 试剂与仪器

组织基因组 DNA 提取试剂盒(TIANamp Genomic DNA KIT),购自 TIANGEN 公司;2 × TaqMan Genotyping Master Mix(货号 4371355),购自美国 Applied Biosystems 公司;荧光定量 PCR 使用 Stratagene Mx3000p PCR 仪;DNA 浓度测定用德国 Eppendorf 公司 Biophotometer 紫外分光光度计。

1.3 试验方法

1.3.1 猪耳组织 DNA 的提取 按 DNA 提取试剂盒的说明步骤提取基因组 DNA,用 Biophotometer 测定样品的基因组 DNA 浓度后置于 -20 °C 保存备用。

1.3.2 引物的设计与合成 从 Ensembl 数据库中下载猪 *KIT* 基因全序列,与 GenBank 中 *KIT* 的 mRNA 序列进行比对,区分 *KIT* 基因的 21 个外显子,用 Primer Express 2.0 于外显子 2 中设计目的基因 *KIT* 的 TaqMan-MGB 探针和引物。引物序列和探针序列(5' → 3')分别为:KITF:GTCAGT-GCTGGCGATGAGATTAG, KITR: ACCCAGGG TCTCAAAGTCCAT; KITPro: FAM-CCGATC-CAGGATCTGT-NFQ-MGB。选取猪单拷贝基因

ESR(雌激素受体)作为定量试验的内标基因,探针和引物参见文献[6]。以上探针、引物均由上海基康生物技术有限公司合成。

1.3.3 实时定量 PCR 体系及反应条件 扩增反应总体积 20 μL 包含 2 × TaqMan Genotyping Master Mix 10 μL, *KIT* 引物 500 nmol · L⁻¹、MGB 探针 250 nmol · L⁻¹(或 *ESR* 引物 900 nmol · L⁻¹、TaqMan 探针 250 nmol · L⁻¹),模板 DNA 4 μL。PCR 扩增条件为:95 °C 预变性 10 min;95 °C 变性 15 s,60 °C 退火/延伸 60 s,40 个循环。

1.3.4 预试验(TaqMan-MGB 探针特异性和稳定性分析) 选取原种杜洛克样品一枚,以一系列 2 倍梯度浓度稀释的 DNA(稀释梯度 2⁰ ~ 2⁴ ng · μL⁻¹)为模板对 *KIT* 基因进行扩增,浓度设置 3 次重复,并设立无模板对照(No Template Control,NTC)。

1.3.5 标准曲线的建立 试验采用单一扩增的方法,即目的基因(*KIT*)和内标基因(*ESR*)的扩增反应在分开的两枚反应管中进行。选取一枚原种杜洛克样品作为标准样品,以一系列 2 倍梯度浓度稀释的 DNA(稀释梯度 2⁰ ~ 2⁴ ng · μL⁻¹)为模板对目的基因(*KIT*)和内标基因(*ESR*)进行定量扩增,每个浓度 3 个重复并设 NTC 对照。以 DNA 相对浓度为横坐标,以 Ct 值为纵坐标,分别绘制 *KIT* 和 *ESR* 基因的标准曲线。

1.3.6 待测样品 *KIT* 基因拷贝数的测定 分别对 50 枚未知样品(体系内终浓度分别为 1 ng · μL⁻¹)的 *KIT* 和 *ESR* 基因进行定量扩增。每次反应,包含一枚标准品样品、一枚 NTC 对照。每个样品仍设置 3 个重复。

当目的基因和参考基因的 PCR 扩增效率近似一致时,待测样品 *KIT* 基因的拷贝数(相对值)为 2^{ΔΔCt},其中,ΔΔCt = ΔCt_{sam} - ΔCt_{cal},ΔCt_{sam} 为待测样品 *KIT* 基因的 Ct 值减去 *ESR* 内标基因的 Ct 值,ΔCt_{cal} 为标准样品 *KIT* 基因的 Ct 值减去 *ESR* 内标基因的 Ct 值^[10]。

1.4 统计学分析

采用 R 软件(R 2.10.0)对数据进行统计分析和统计图绘制,由于 1.3.6 中计算的拷贝数相对值呈连续分布的实数,为便于合理分类,本研究用聚类分析的方法对 50 个未知样品 *KIT* 基因的拷贝数进行归类。

2 结 果

2.1 预试验 TaqMan-MGB 探针检测效果分析

原始 DNA 浓度 $16 \text{ ng} \cdot \mu\text{L}^{-1}$ 按 2 倍稀释至 $1 \text{ ng} \cdot \mu\text{L}^{-1}$ 得到 5 种不同浓度的系列工作液, 经过 TaqMan-MGB 探针荧光实时定量 PCR 检测得到的扩增曲线见图 1。由图 1 可见, 不同浓度组间 *KIT* 基因的扩增曲线呈现间距相等的平行曲线, 每个浓度的 Ct 值相差约 1 个循环, 表明该方法检测基因的拷贝数具有良好的线性范围。进一步的分析表明, 标准曲线的直线相关系数 $r^2 = 0.999$, 说明线性关系非常好。扩增产物测序结果与扩增的目的片段完全吻合, 表明该方法的特异性好。

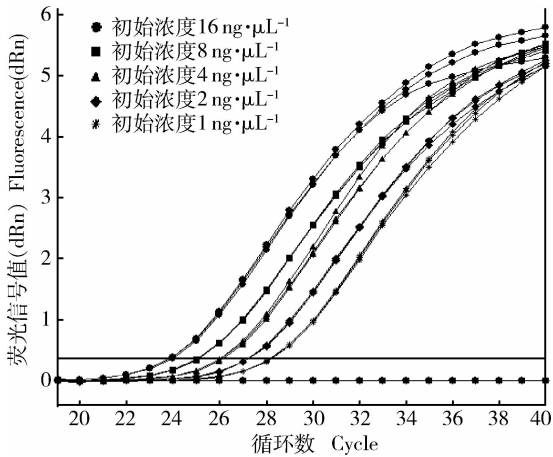


图 1 2 倍系列稀释 DNA 样品 *KIT* 基因的扩增曲线

Fig. 1 Amplification curve of *KIT* obtained with 2-fold serial dilutions of DNA samples

表 1 列出了每个 DNA 浓度下 3 次重复测定的 *KIT* 基因 Ct 值的平均数 (\bar{C}_t)、标准差 (SD) 和变异系数 (C. V)。本试验测定的 Ct 值变异系数的变异范围从 0.13% ~ 0.26%, 均在 1% 的合理范围内, 表明 TaqMan-MGB 实时定量 PCR 方法具有很好的稳定性和可重复性。

2.2 标准曲线

以 *KIT* 和 *ESR* 各浓度的 Ct 平均值为纵坐标, 以 DNA 相对浓度的 log 值为横坐标绘制标准曲线 (图 2)。由图 2 可知, *KIT* 和 *ESR* 对应的 2 个扩增标准曲线的斜率之差为 0.061, 小于 0.1, 扩增效率分别为 97.0% 和 99.4%, 相差 2.4% (小于 5% 即认为一致), 这说明目标基因和内标基因的扩增效率近

似一致, 可以通过 $2^{-\Delta\Delta C_t}$ 方法对 *KIT* 进行相对定量^[10]。

表 1 TaqMan-MGB 实时定量 PCR 检测 *KIT* 基因拷贝数的稳定性

Table 1 Stability of TaqMan-MGB Real-time quantitative PCR

项目 Item	DNA 浓度 / ($\text{ng} \cdot \mu\text{L}^{-1}$) DNA concentration				
	1	2	4	8	16
\bar{C}_t	28.20	27.25	26.07	25.08	23.90
SD	0.04	0.04	0.07	0.06	0.04
C. V / %	0.14	0.13	0.26	0.24	0.17

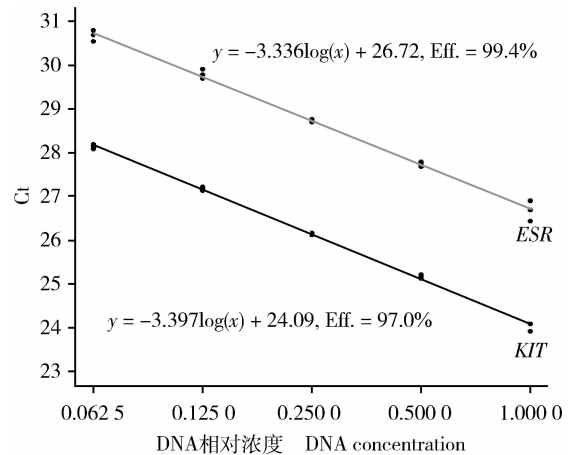


图 2 标准品 *KIT* 和 *ESR* 基因的标准曲线

Fig. 2 Standard curves for *KIT* and *ESR* of calibrator

2.3 未知样品 *KIT* 基因的定量检测

对 50 头猪的基因拷贝数进行了定量检测, 记录所得 Ct 值。对样品 (N=50) 的 $2^{-\Delta\Delta C_t}$ 值进行聚类分析, 区分未知样品 *KIT* 基因的拷贝数 (图 3)。当类间距离小于 0.4 时, 可将 50 个样品归入 5 个不同的类, 即分别对应 2、3、4、5、6 个拷贝数。图 4 表示了不同拷贝数的样品对应的 $2^{-\Delta\Delta C_t}$ 值, 2、3、4、5、6 个拷贝重复数对应的 $2^{-\Delta\Delta C_t}$ 平均值 ± 标准误分别是: 0.89 ± 0.04 , 1.28 ± 0.03 , 1.71 ± 0.02 , 2.13 ± 0.04 , 3.25 ± 0.05 , 各组均数间差异极显著 ($P < 0.001$)。6 头杜洛克均为 2 拷贝, 大白猪为 5 或 6 个拷贝, 白色杜洛克除了 1 头为 2 拷贝外, 其余 37 头为 3~6 拷贝。

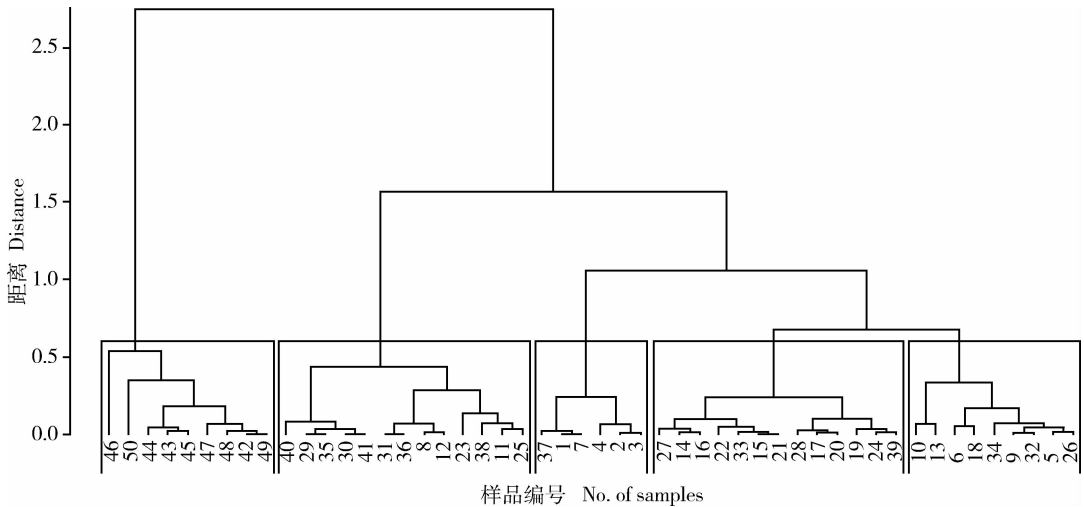


图3 待测样品 $2^{-\Delta\Delta C_t}$ 值的系统聚类图 (N=50)

Fig. 3 Hierarchical clustering results of unknown samples by $2^{-\Delta\Delta C_t}$ (N=50)

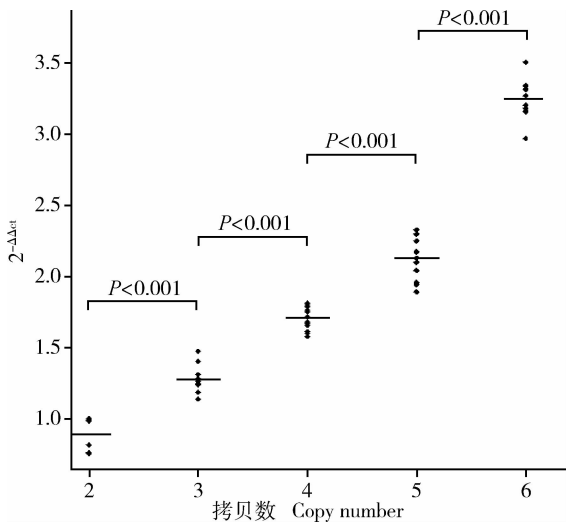


图4 *KIT* 基因拷贝重复数的相对定量值 $2^{-\Delta\Delta C_t}$ (N=50)

Fig. 4 Relative quantification ($2^{-\Delta\Delta C_t}$) of copy numbers of *KIT* (N=50)

3 讨论

TaqMan-MGB 探针由于采用特殊的设计,特异性、灵敏度、稳定性比普通探针大大提高^[11]。从本研究结果来看,TaqMan-MGB 探针得到的扩增曲线特异性和稳定性都很好,扩增效率高,为 *KIT* 基因拷贝重复数的定量测定提供了很好的基础。本研究建立的以 TaqMan-MGB 探针为基础的荧光定量 PCR 检测 *KIT* 基因拷贝重复数的方法,还具有操作简便、快捷(3~5 h 可完成测定)的特点,不需昂贵设备,在一般实验室即可完成。

尽管我们知道衡量样品 *KIT* 拷贝重复数的值- $2^{-\Delta\Delta C_t}$ 与拷贝数呈正相关,但对未知样品检测得到的 $2^{-\Delta\Delta C_t}$ 值呈现连续变异,究竟每个样品含几个拷贝需要合理分组。为避免人为分组造成的偏差,本研究采用系统聚类分析方法根据样品间数据差异的大小(即距离)将样品进行聚类分组,因为相同拷贝数样品间的 $2^{-\Delta\Delta C_t}$ 值会更接近。聚类分组结果表明,不同拷贝数组间的 $2^{-\Delta\Delta C_t}$ 平均值有显著差异且标准误差小,说明分类较为合理。除了拷贝数最高的一组外,其余 4 组 $2^{-\Delta\Delta C_t}$ 的比值近似为 2 : 3 : 4 : 5,依次对应 2、3、4、5 个拷贝数。拷贝数较高组的 $2^{-\Delta\Delta C_t}$ 值偏高,可能意味着猪群中有 6 个以上拷贝数的存在,这与 Pielberg 等^[5] 的检测结果有相似之处。各组间的 $2^{-\Delta\Delta C_t}$ 无重叠现象,试验误差小,检测效果好于文献^[5] 中使用的普通 TaqMan 探针实时定量 PCR 方法。

6 头杜洛克猪的 *KIT* 均为 2 拷贝,大白猪 5 或 6 个拷贝各 3 头。白色杜洛克猪除了 1 头为 2 拷贝外,其余 37 头为 3~6 拷贝。白色杜洛克系普通杜洛克与大约克猪杂交选育的被毛白色的种群,*KIT* 基因重复拷贝数的分布符合预期值。

4 结论

本研究建立了 TaqMan-MGB 实时荧光定量 PCR 检测 *KIT* 基因拷贝重复数变异的方法,该方法可以直接对产物进行定量,具有良好的稳定性和特异性,而且操作简单、安全,自动化程度高,不易产生污染,不需昂贵设备,适合在普通实验室检测。

参考文献:

- [1] JOHANSSON M, ELLEGREN H, ANDERSSON L, et al. The gene for dominant white color in the pig is closely linked to *ALB* and *PDGFRA* on chromosome 8 [J]. *Genomics*, 1992, 14(4):965-969.
- [2] JOHANSSON M, CHAUDHARY R, ANDERSSON L, et al. Pigs with the dominant white coat color phenotype carry a duplication of the *KIT* gene encoding the mast/stem cell growth factor receptor [J]. *Mamm Genome*, 1996, 7(11):822-830.
- [3] WEHRLE-HALLER B, WESTON J. Receptor tyrosinekinase dependent neural crest migration in response to differentially localized growth factors [J]. *Bioessays*, 1997, 19(4): 337-345.
- [4] MARKLUND S, KIJAS J, RODRIGUEZ-MARTINEZ H, et al. Molecular basis for the dominant white phenotype in the domestic pig [J]. *Genome Res*, 1998, 8(8):826-833.
- [5] PIELBERG G, OLSSON C, ANDERSSON L, et al. Unexpectedly high allelic diversity at the *KIT* locus causing dominant white color in the domestic pig [J]. *Genetics*, 2002, 160(1): 305-311.
- [6] GIUFFRA E, EVANS G, ANDERSSON L, et al. The Belt mutation in pigs is an allele at the Dominant white (*I/KIT*) locus [J]. *Mamm Genome*, 1999, 10(12): 1132-1136.
- [7] PIELBERG G, DAY A, PLASTOW G, et al. A sensitive method for detecting variation in copy numbers of duplicated genes [J]. *Genome Res*, 2003, 13(9): 2171-2177.
- [8] SEO B, PARK E, AHN S, et al. An accurate method for quantifying and analyzing copy number variation in porcine *KIT* by an oligonucleotide ligation assay [J]. *BMC Genetics*, 2007, 8: 81.
- [9] HEID C, STEVENS J, LIVAK K, et al. Real-time quantitative PCR [J]. *Genome Res*, 1996, 6(10): 986-994.
- [10] LIVAK K, SCHMITTGEN T. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta Ct}$ Method [J]. *Methods*, 2001, 25(4): 402-408.
- [11] KUMAR S, REED M, SCHWEITZER B, et al. Solution structure of a highly stable DNA duplex conjugated to a minor groove binder [J]. *Nucleic Acids Res*, 1998, 26(3): 831-838.

(编辑 郭云雁)