

猪 *TAF7* 基因的克隆、染色体定位和组织表达谱分析

张兴举, 唐中林, 王志伟, 杨述林, 牟玉莲, 崔文涛, 马月辉, 储明星, 李奎

(中国农业科学院北京畜牧兽医研究所/中国农业科学院家养动物遗传资源与种质创新重点开放实验室, 北京 100193)

摘要: 【目的】通过对猪 *TAF7* 基因初步的研究, 为猪分子遗传育种提供基础分子生物学信息, 为猪的遗传育种提供分子标记。【方法】以五指山猪为研究对象, 克隆 *TAF7* 基因, 分析该基因结构特点, 然后利用 IMpRH (the INRA-University of Minnesota porcine radiation hybrid, 法国农业科学院-明尼苏达大学的辐射杂种克隆板) 分析该基因在猪染色体上定位信息, 利用半定量 RT-PCR 方法, 分析该基因在成年五指山猪 16 个不同组织(心脏、背肌、淋巴、脾脏、肝脏、肾脏、肺脏、子宫、睾丸、胃、小肠、大肠、卵巢、胸腺、脑、脂肪)的表达谱信息。【结果】克隆得到长 1 701 bp 的五指山猪 *TAF7* 基因序列, 其中包括 1 050 bp 完整 CDS (Coding Sequence, 编码序列) 区域, 分析表明其编码含 349 个氨基酸的蛋白质。利用 IMpRH 分析结果表明, 猪 *TAF7* 基因与分子标记 SW1879 和 *IL4* (interleukin-4, 白细胞介素-4) 紧密连锁, LOD (Limit of Detection, 检测极限) 值分别为 6.69 和 6.15。组织表达谱分析结果显示该基因在大多数组织中均有表达, 其中在睾丸表达量较高, 而在心脏、背肌中表达很低。【结论】猪 *TAF7* 基因 5'UTR 中有一个短的内含子, 而在该基因 CDS 区域没有内含子。猪 *TAF7* 蛋白序列与人 *TAF7* 蛋白序列的相似性较高, 二者在生物系统发育树中的距离最接近。猪 *TAF7* 基因在大多数组织中均表达, 在猪睾丸中表达量较高, 证实它调节的基因转录具有广泛性。

关键词: 猪 *TAF7* 基因; 克隆; 基因定位; 组织表达谱

Cloning, Mapping and Tissue Expression Profile of Porcine *TAF7* Gene

ZHANG Xing-ju, TANG Zhong-lin, WANG Zhi-wei, YANG Shu-lin, MU Yu-lian, CUI Wen-tao, MA Yue-hui, CHU Ming-xing, LI Kui

(The Key Laboratory of Domestic Animal Genetic Resources and Germplasm Innovation of CAAS, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193)

Abstract: 【Objective】 The objective is to illuminate the molecular features and the expression profile of the porcine *TAF7* (TBP-associated factor; TBP is a TATA-binding protein) gene. As one component of the transcriptional pre-initiation complex, the protein *TAF7* plays an important role during regulation of the gene expression. 【Method】 The porcine *TAF7* gene from the Wuzhishan pig sample was cloned and sequenced, and then its molecular features were analyzed. Mapping of the porcine *TAF7* gene to porcine chromosome was made with the IMpRH panel. The mRNA distribution profile of the porcine *TAF7* gene in sixteen adult Wuzhishan pig tissues (heart, dorsal muscles, lymph nodes, spleen, liver, kidney, lung, uterus, testicle, stomach, small and large intestine, ovaries, thymus gland, brain and fat) were examined by semi-quantitative RT-PCR. 【Result】 A sequence with a length of 1 701 bp was cloned, it contained the CDS (coding sequence) region of 1 050 bp of the porcine *TAF7* gene. Analysis of the network software GENSCAN showed that the porcine *TAF7* gene encoded a protein composed of 349 AAs. By the IMpRH (the INRA-University of Minnesota porcine radiation hybrid) panel, porcine *TAF7* gene was assigned to porcine chromosome 2, associated with SW1879 and *IL4* (interleukin-4) closely, the LODs (Limit of Detection) were 6.69 and 6.15, respectively. The mRNA distribution of the porcine *TAF7* gene was that it was expressed almost in all tissues, especially highly expressed in testicle and could hardly be examined in heart and dorsal muscles. 【Conclusion】 The porcine *TAF7* gene has a small intron in the 5'UTR, but the

收稿日期: 2008-02-28; 接受日期: 2008-05-13

基金项目: 国家自然科学基金 (30571300), 中国农业科学院北京畜牧兽医研究所“家畜种质资源研究和创新”科技创新团队资助

作者简介: 张兴举 (1980-), 男, 河北巨鹿人, 硕士研究生, 研究方向为动物分子生物学。E-mail: xingju_0128@163.com。通信作者李奎 (1963-), 男, 河北松滋人, 教授, 博士, 研究方向为动物遗传育种与繁殖。Tel: 010-62813822; Fax: 010-62813822; E-mail: likuihau@yahoo.com

CDS is intronless. The sequence of the porcine protein TAF7 shares high similarity with its homolog from human, and they both are closed in the Phylogenetic tree. The porcine *TAF7* gene can be detected almost in all tissues, which shows us its extensive functions for transcription. Furthermore, the phenomenon of highly expressed in testicle needs further research.

Key words: porcine *TAF7* gene; cloning; mapping; expression profile

0 引言

【研究意义】TAF7 蛋白是很多基因转录的初始转录复合物成分之一，它在基因表达调控中起到非常重要的作用，克隆猪 *TAF7* 基因，明确在染色体上的精细定位及其遗传连锁的分子标记，对猪遗传学研究和猪分子育种学研究具有重要意义。【前人研究进展】TF II D (transcription factor II D, 转录因子 II D) 在真核生物基因转录的启动过程中起到中枢性作用^[1]。TF II D 复合物由 TBP (TATA-binding protein, TATA 框结合蛋白) 与 TRF (TBP-related factor, TBP 关联蛋白) 和 TAFs (TBP-associated factors, TBP 相关蛋白) 组成^[2-6]，这 3 类转录因子均具备各自独特的生物学功能，其中 TAF 是一个具备转录起始识别和调控的独立的蛋白家族。近年来，越来越多的 TAF 家族成员被鉴定出来，研究发现 TAF 家族成员在氨基酸序列上具有很高的相似性，而且每一个成员之间在不同物种间具有较高的保守性^[7]。在 TF II D 复合物中，TAF1 作为“支架”蛋白同时具有 DNA 和其它转录因子结合的活性，普遍存在于转录起始复合物中。TAF7 被称为转录起始复合物中的抑制因子，它直接与 TAF1 的乙酰转移酶 (AT) 亚基结合，通过抑制 TAF1 的乙酰基转移酶活性而达到抑制基因表达、调节基因转录的功能^[8]。近年来，关于 TAF7 蛋白的功能及其分子作用机制的研究越来越多。一般认为，TAF 通过两条途径调控各种生物学现象，一是与其它共激活因子互动，以共同调控目的基因的转录；二是直接结合启动子元件，以控制目的基因的转录。Lavigne 研究发现，维生素 D₃、甲状腺激素受体和视黄醇类 X 受体衍生物均可与 TAF7 蛋白互作，促使 TAF7 转录活性提高 5 倍^[9]；c-Jun 作为 TAF7 的共激活因子，通过与 TAF7 结合改变 TAF7 结构来调节某些基因的表达，达到对细胞信号应答的目的^[10]。在精子发生过程中，从精母细胞发育到粗线期开始，TAF7 从 TF II D 复合物中脱离并在细胞核内逐渐消失，TAF1 的乙酰基转移酶活性恢复，同时 TAF7L 因子取代 TAF7 的位置，TAF7L 与 TBP 和 TAF 家族成员结合并负责精子发生过程中一系列蛋白的转录^[11-12]。【本研究切入点】TAF7 作

为基因转录起始因子之一，在基因表达调控中发挥广泛的、重要的生物学作用。目前已克隆了酵母、果蝇、人、鼠、牛等 11 个物种的 *TAF7* 基因，对其功能也有了一定的了解。但对于 TAF7 蛋白的生物学作用未了解透彻，需要进一步研究，而且未见有关猪 *TAF7* 基因的研究报道，因此，本试验以五指山猪为研究对象，开展猪 *TAF7* 基因的初步研究。【拟解决的关键问题】克隆猪 *TAF7* 基因，确定该基因在猪染色体上的精细定位，分析其遗传连锁信息，分析该基因在成年五指山猪 16 个组织中的表达情况。

1 材料与方法

1.1 基因克隆

本试验所用 30 个 DNA 样品分别来自五指山小型猪、巴马小型猪、贵州小型猪、莱芜猪、通城猪和长白猪耳组织，每种猪采 5 个个体的样品。本试验使用五指山猪睾丸组织 cDNA 样品，RNA 提取按照 Trizol 试剂盒 (Life Technologies, Grand Island, NE, USA) 操作说明进行，经 RT-PCR 获得 cDNA，置于 -20℃ 保存。

根据 NCBI 网站数据库中提供的人 *TAF7* 基因序列钓取猪 EST，将所得猪 EST 经 DNASTar 软件拼接得到 contig 序列，按照此 contig 序列利用 Primer Premier 5.0 设计 PCR 引物，引物信息见表 1。

用琼脂糖凝胶 DNA 回收试剂盒 (TianGen) 回收纯化 PCR 扩增产物，将目的片段与 pGEM-T easy 载体连接，16℃ 连接 6 h。然后把连接产物转入 DH5α 感受态细胞，通过蓝白斑筛选，挑取阳性克隆。经 PCR 检测后，将阳性菌液并送至北京英俊生物技术公司测序。

1.2 染色体定位

利用法国农业科学院明尼苏达大学提供的辐射杂种克隆板 (含 118 个杂种细胞系 DNA 模版) 对猪 *TAF7* 基因进行染色体定位。然而，应用辐射杂种克隆板对染色体进行精细定位对于 PCR 引物的要求比较高，本试验中设计的引物 (TMSF/TMSR) 如表所示。确定使用该引物扩增在猪模板中呈阳性，而在仓鼠模板中呈阴性。同时，PCR 产物片段具有如下特点：(1)

表 猪 *TAF7* 基因克隆、定位和表达谱分析的引物Table Primers used for cloning, mapping and express profile analyzing of porcine *TAF7* gene

引物名称 Primers	引物序列(5'-3') The sequence of primer 5'-3'	产物大小(bp) The length of the product	退火 Tm (°C) Renaturation temperature	延伸时间(s) Time for extension
T1F	TTTGCTGCGAGTGGCTTCC	662	64	30
T1R	ACGCCATCGGGGTGTAA			
T2F	CAGTTTATCTTACGCCTACCT	1153	64	40
T2R	GTCGATATGACAAACCCAAT			
TT1F	TTTGCTGCGAGTGGCTTCC	867	58	35
TT1R	TCTTGCTTGCTTTGGGATC			
TT2F	AGCTGCTACTGCTGATCCC	644	58	30
TT2R	GTTCATCCAGCACGGCCT			
TT3F	CAGATTTCAGGCCGTGCT	248	60	20
TT3R	GTCGATATGACAAACCCAAT			
TEF	TTTGCTGCGAGTGGCTTCC	662	65	30
TER	ACGCCATCGGGGTGTAA			
TMSF	TGACTCACTGGAACACGATG	280	61	20
TMSR	ATCCTCCTGTCGCTTTGC			
βF	GGACTTCGAGCAGGAGATGG	233	60	20
βR	GCACCGTGTGGCGTAGGG			

引物 T1F/T1R 和 T2F/T2R 用于 mRNA 扩增; 引物 TT1F/TT1R、TT2F/TT2R 和 TT3F/TT3R 用于基因组扩增; 引物 TEF/TER 用于表达谱; 引物 TMSF/TMSR 用于染色体定位和 SNP 扩增; 引物 βF/βR 用于表达谱内参扩增
The primers T1F/T1R and T2F/T2R were used for amplifying mRNA of porcine *TAF7*, TT1F/TT1R, TT2F/TT2R and TT3F/TT3R for amplifying the introns, TEF/TER for analyzing mRNA tissue distribution, TMSF/TMSR for mapping; βF/βR as control in tissue distribution

产物片段长 280 bp (小于 300), 容易扩增得到; (2) PCR 反应条件下产物产量很大, 容易检测; (3) PCR 产物纯度高, 电泳条带单一。

本试验 PCR 采用 10 μl PCR 扩增体系, 其中含 25 ng 模板 DNA、2.5 mmol·L⁻¹ MgCl₂、1×PCR 缓冲液、0.35 U Taq DNA 聚合酶、dNTP 50 μmol·L⁻¹、上、下游引物各 0.25 μmol·L⁻¹。每次 PCR 都包含猪 DNA、仓鼠 DNA 和不含任何 DNA 模板的 ddH₂O 作为 3 个对照样。PCR 反应条件是: 94°C 变性 5 min, 30 个循环 (94°C 30 s, 61°C 退火 30 s, 72°C 延伸 20 s), 最后 72°C 延伸 5 min。PCR 产物用 1.5% 琼脂糖凝胶电泳, EB 染色。PCR 结果提交到法国农业科学院-明尼苏达大学的在线分析工具 <http://www.toulouse.inra.fr/lgc/pig/compare/HSA.htm> 进行分析。

1.3 组织表达谱

采集成年五指山猪 16 种组织 (心脏、背肌、淋巴、脾脏、肝脏、肾脏、肺脏、子宫、睾丸、胃、小肠、大肠、卵巢、胸腺、脑、脂肪) 迅速放到液氮中冷冻, 然后于 -80°C 保存。采用 Trizol 试剂盒提取组织总 RNA, 并反转录为 cDNA 样品, 于 -20°C 保存。根据

拼接猪 contig 序列, 利用 Primer Premier 5.0 软件设计用于组织表达谱分析的 PCR 引物 TEF/TER, 引物信息见表。因为 TAF 家族成员之间的核酸序列有较高的相似性, 为了保证扩增的特异性, 首先, 在设计引物前分析了 TAF 家族成员基因序列的同源性, 将引物设计在 *TAF7* 基因 cDNA 序列特异位点; 其次设定稍高一些的退火温度 (65°C), 进一步避免非特异性扩增; 最后以 β-actin 作为内参, 将扩增循环数降低为 28, 将模板浓度调整一致。PCR 条件参考“表 1”引物 TEF/TER 信息, 其他扩增体系和扩增条件参照 1.2 中所述。PCR 产物于 1.0% 琼脂糖凝胶中电泳, 以 EB 染色, 并在紫外线下检测。

2 结果与分析

2.1 基因序列特征

根据人 *TAF7* 基因序列通过 BLAST 软件钩取猪 EST 序列, 把这些 EST 经 DNASTar 软件拼接后获得了一条长 2 450 bp 的猪 contig 序列。根据该 contig 序列设计引物 T1F/T1R 和 T2F/T2R, PCR 扩增片段在北京英俊生物公司测序, 测序结果应用 DNASTar 软件进

拼接,得到长 1 701 bp 的脱氧核苷酸序列。与人 *TAF7* 基因序列比对结果表明,此序列含猪 *TAF7* 基因上 506 bp 的 5'UTR、145 bp 的 3'UTR 以及中间 1 050 bp 的 CDS 区域。通过网络软件 GENSCAN (<http://genes.mit.edu/GENSCAN.html>) 分析得到猪 *TAF7* 基因编码的猪 *TAF7* 蛋白含 349 个氨基酸序列(图 1), 分子量为 4 0217.90, 等电点为 4.8331(<http://weblab.cbi.pku.edu.cn/program.inputForm.do?program=pepstats>)。利

用 NCBI 网络软件 rpsblast (<http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi?cmd=rps>) 对猪 *TAF7* 蛋白进行保守结构域预测(图 2), 发现含有人转录因子 TAFII55(人 *TAF7* 另一个名字)保守结构域。已知在人 *TAF7* 蛋白中, 该结构域的功能是参与基因转录起始, 有研究表明 TAFII55 能与另一转录因子 TAFII250(*TAF1* 另一个名字)结合并抑制其乙酰活性, 以达到抑制基因转录的结果^[13-14]。

```

TTTGCTGCGAGTGGCTTCCTTTCTTTCCGTTCCGTTCCGGTCCGGAGAGACTGGATTTAATAATCGTCACTGGATTGTAAGTACCCGAGGCGAAGAGAGCTCACTACGCCCCGCTCTTGAATAT 120
CTTTGTTCCGGGAGAGTTTGTGCATTGAGAGTTTCTGCTGCGCCTTCTCTTCTGTGCCAAAAATCCTCCGTAGGGGCGGCCATTGCTTTCCCTGCTAGCTAGCTCTCCCGGTGGCCTG 240
CAGCGGCATCTAGAGGTTGGAAC TTGGCATTGCGAGCTAATAGATTTAAAAGGACTTAAGTGTTTGAACCACTGGTATTGAGTCAACAGCTCTTTAGAAAGACTTGGGCTTAAATTCACA 360
ATAGACGCTTAGGTGCTTTTCTACTTAGAGCTTTTGTACTTAAAGACTTCAATTTATGACGAACTAGTCTTCTGTTGACATAGCCGGAACCTTTTAGTCCAAACAAGTGGTTGATAGAG 480
AAAAAGTGAGTCCACAGAGAAATAAAGATGAGTAAGAGCAAGAGCGATGCTCCTCATGAACCTGGAGAGCCAGTTTATCTTACGCC TACTCCGGAATATGCTCTACGGTGAGCGGGCAG 600
Met Ser Lys Ser Lys Asp Asp Ala Pro His Glu Leu Glu Ser Gln Phe Ile Leu Arg Leu Pro Pro Glu Tyr Ala Ser Thr Val Arg Arg Ala 31
TACAGTCTGGCCATGTCAACCTGAAGGACAGACTGACAATTGAGTTACACCCGATGGCGTCATGGAAATCGTCAGAGTGGACCGGGTCCACTGGCCTCAAAAATGGTAGATCTGCCCT 720
Val Gln Ser Gly His Val Asn Leu Lys Asp Arg Leu Thr Ile Glu Leu His Pro Asp Gly Arg His Gly Ile Val Arg Val Asp Arg Val Pro Leu Ala Ser Lys Leu Val Asp Leu Pro 71
GTGTTATGGAGAGTTTGAACCAATTGTAAGAAAACCTTTTACAAGCAGCTGATATCTGTCAGATGCTCGTCTCTACAGTTGATGGTATCTCTATCCCTGTGGGAAACCAAGCTG 840
Cys Val Met Glu Ser Leu Lys Thr Ile Asp Lys Lys Thr Phe Tyr Lys Thr Ala Asp Ile Cys Gln Met Leu Val Ser Thr Val Asp Gly Asp Leu Tyr Pro Val Glu Glu Pro Ala 111
CTACTGCTGATCCCAAGCAAGCAAGAAAGGATAAAGGACAAAGAGAAAGAGTTTATATGGAACCTGGAATTAAGTCTGCTCTAAAAATGTCGAAAGAGAGAGGTTCCGAAAGACAG 960
Ala Thr Ala Asp Pro Lys Ala Ser Lys Lys Asp Lys Asp Lys Glu Lys Lys Phe Ile Trp Asn His Gly Ile Thr Leu Pro Leu Lys Asn Val Arg Lys Arg Phe Arg Lys Thr 151
CAAAGAAGAAGTACATTGAGTCTCCGGAATGTGAAAAAAGAAGTGAAGCGCTTGTGAGCACAGATGCCGAAGCTGTCACTACCCGCTGGGAAAT AATTGCTGAAGATGAAACAAAAGAAA 1080
Ala Lys Lys Lys Tyr Ile Glu Ser Pro Asp Val Glu Lys Glu Val Lys Arg Leu Leu Ser Thr Asp Ala Glu Ala Val Ser Thr Arg Trp Glu Ile Ile Ala Glu Asp Glu Thr Lys Glu 191
CAGAAAATCAAGGCCTTGATATCTCTTCCAGGAATGTCTGGCCACAGGCGAGGCCATGACTACTGGAACACGATGAGCTTCGGGAGATATTCAATGACCTCAGCAGCAGCAGTGAGG 1200
Thr Glu Asn Gln Gly Leu Asp Ile Ser Ser Pro Gly Met Ser Gly His Arg Gln Gly His Asp Ser Leu Glu His Asp Glu Leu Arg Glu Ile Phe Asn Asp Leu Ser Ser Ser Glu 231
ATGAAGATGAGACACAGCATCAAGATGAAGAATATAAACATCATCGACCGGAGGAAAGACCTGGAAGGCGAGCTACAGGACAAGCTCAGTGAATCAGATGAACAGCACCAGAAAATG 1320
Asp Glu Asp Glu Thr Gln His Gln Asp Glu Glu Asp Ile Asn Ile Ile Asp Thr Glu Glu Asp Leu Glu Arg Gln Leu Glu Asp Lys Leu Ser Glu Ser Asp Glu Gln His Gln Glu Asn 271
AGGGAACCAATCAAGTGGTGGAAATTCAGAAACAGATTGATAACATGAAAGGCAAGCTCCAGAGACCCAGGACAGGGCAAAGCGACAGGAGGATCTGATCATGAAAGGTGAAAATC 1440
Glu Gly Thr Asn Gln Leu Val Met Gly Ile Gln Lys Gln Ile Asp Asn Met Lys Gly Lys Leu Gln Thr Thr Gln Asp Arg Ala Lys Arg Gln Glu Asp Leu Ile Met Lys Val Glu Asn 311
TGGCTCTCAAGAACAGATTTCAAGCCGTGCTGGATGAACTGAAACAAAAGGAAAGACCAGAAAAAGGAGCAGCTCAGCTCTTTGCAGGAGGAGCTAGAATCACTCTGGAGAAATGAGGAG 1560
Leu Ala Leu Lys Asn Arg Phe Gln Ala Val Leu Asp Glu Leu Lys Gln Lys Glu Asp Arg Glu Lys Glu Gln Leu Ser Leu Glu Glu Leu Glu Ser Leu Leu Glu Lys * 349
AGCTTTGATACTTAATTTTCTTCTCAGAGTAGTCGGCATTGAAAGAAAATTTGGTTTCATCACTGGACTAGGTACTATATAACTTTGCAGTAGTGGCCATTTTCTTCACTGCTCTC 1680
TATTGGGTTTGTCCATATCGAC 1701

```

图 1 猪 *TAF7* 基因及其编码蛋白质序列

Fig. 1 The cDNA and protein sequences of porcine *TAF7* gene



图 2 猪 *TAF7* 蛋白序列保守结构域预测

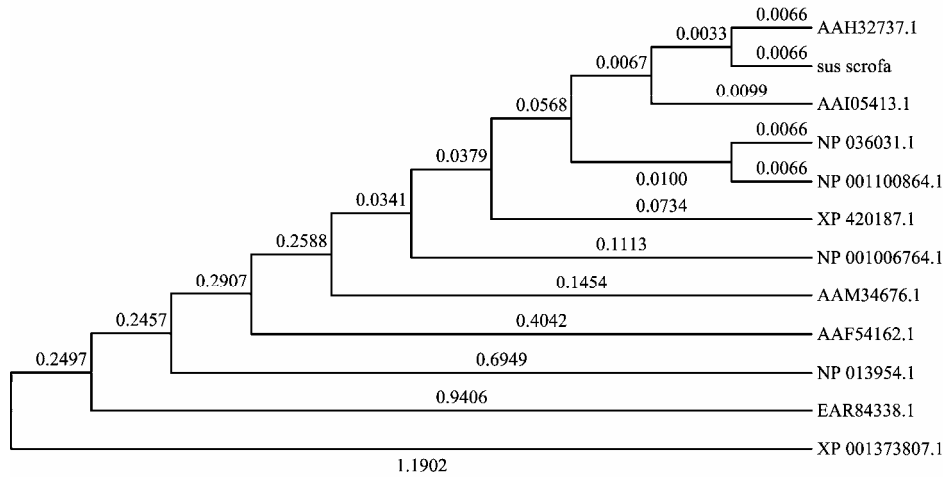
Fig. 2 Prediction of functional domain of porcine *TAF7*

TAF 家族成员在物种间的保守性很高, 本试验对此做了进一步的验证。将 NCBI 公布的所有的 12 个物种 *TAF7* 蛋白序列及猪 *TAF7* 蛋白序列输入 MEGA3.1 软件, 构建了系统发育树(图 3)。图示人和猪的 *TAF7* 蛋白序列进化关系最为紧密, 说明以往对人 *TAF7* 蛋白研究结果可以指导对猪 *TAF7* 蛋白结构和功能的研究。

本试验对猪 *TAF7* 基因结构作了进一步分析, 通过 TT1F/TT1R、TT2F/TT2R 和 TT3F/TT3R 三对引物扩增五指山猪基因组 DNA, 得到猪 *TAF7* 基因序列全长(包括内含子和外显子)。3 对引物的 PCR 产物序

列拼接后得到 1 704 bp 的 DNA 序列, 与已克隆得到的猪 cDNA 序列比对, 发现猪 *TAF7* 基因 cDNA 编码区没有内含子, 而在该基因 5'UTR(本试验提交序列中第 298 与 299 碱基之间)有一个长度仅为 3nt 的内含子, 序列为“TAA”(5'到 3'方向), 这与人和鼠 *TAF7* 基因结构(无内含子)略有不同。

本试验测定了 6 个不同猪品种(五指山小型猪、巴马小型猪、贵州小型猪、莱芜猪、通城猪和长白猪)共计 30 只猪的 *TAF7* 基因序列。将这 30 条序列送至北京英俊生物公司测序, 测序结果利用 DNAStar 软件进行比对分析, 发现该基因第 1 256 个碱基发生 G/C



图中“sus scrofa”：本试验预测猪 TAF7 蛋白；NP_013954.1：酿酒酵母；AAH32737.1：人；AAF54162.1：黑腹果蝇；NP_036031.1：小鼠；XP_420187.1：原鸡；NP_001006764.1：热带蟾蜍；AAI05413.1：牛；EAR84338.1：四膜虫；AAM34676.1：斑马鱼；XP_001373807.1：短尾负鼠；NP_001100864.1：褐鼠

“sus scrofa” represents the porcine TAF7 sequence we have predicted; NP_013954.1: Saccharomyces cerevisiae; AAH32737.1: Homo sapiens; AAF54162.1: Drosophila melanogaster; NP_036031.1: Mus musculus; XP_420187.1: Gallus gallus; NP_001006764.1: Xenopus tropicalis; AAI05413.1: Bos Taurus; EAR84338.1: Tetrahymena thermophila SB210; AAM34676.1: Danio rerio; XP_001373807.1: Monodelphis domestica; NP_001100864.1: Rattus norvegicus

图 3 12 个不同物种 TAF7 蛋白系统发育树

Fig. 3 Phylogenetic tree of protein TAF7 of twelve species

突变，该单核苷酸多态性（single nucleotide polymorphism, SNP）是否对猪 TAF7 基因功能有影响尚值得进一步研究。

2.2 染色体定位

利用辐射杂种克隆板扩增效果较好，118 个样品中仅有 4 个样品不能准确判断结果，如图 4-a 所示，

“0”代表阴性，“1”代表阳性，“？”代表无法判断。将该结果提交到法国农业科学院网络软件中，经网络分析后确定猪 TAF7 基因定位于 2 号染色体上，与 IL4 紧密连锁（LOD 值为 6.15），同时与微卫星标记 SW1879 紧密连锁（LOD 值为 6.69），如图 4-b 所示。

Reference map:
 Hawken et al. (118 Hybrids) a

PCR data:

```

0?1010000001000000000?00000000000011101
10000000001001000001001000000?000010011
010000100011100000011010000100010000?1
  
```

Corresponding two points results: b

Order	Chr.	Marker 1 Marker 2	P(br) 1 P(br) 2	Dist 1(Ray) Dist 2(Ray)	LOD 1 LOD 2
1	2	SW1879	0.43	0.57	6.69
		SWR308	0.54	0.79	4.43
2	2	SWR2157	0.65	1.06	2.46
		IL4	0.45	0.60	6.15
3	15	SW1263	0.90	2.29	0.21
		SW120	0.79	1.58	0.86
4	13	S0291	0.83	1.77	0.58
5	2	IL4	0.45	0.60	6.15
		SW1879	0.43	0.57	6.69

图 4 辐射杂种克隆板扩增 118 个样品结果 (a) 及染色体定位分析结果 (b)

Fig. 4 The symbols ‘a’ and ‘b’ are the result of the PCR with the IMpRH panel and the result of mapping

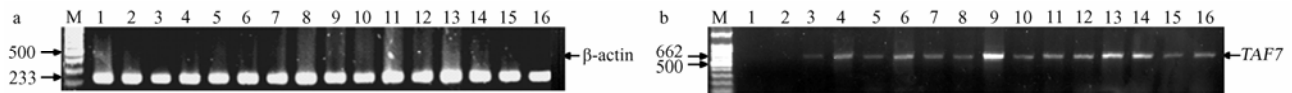
2.3 组织表达谱分析

以 β-actin 为内参，采用半定量 RT-PCR 方法对 TAF7 基因做组织表达谱分析。猪 TAF7 基因在 16 个组织当中的分布如图 5 所示，该基因几乎在所有组织中均有表达，其中在睾丸中表达量最高，说明该 TAF7 作用的广泛性和重要性。另外，研究发现猪 TAF7 基

因在睾丸中表达量较高，而在心脏和背肌中未检测到表达，这一现象需要进一步研究解释。

3 讨论

真核生物的 mRNA 是由 RNA 聚合酶 II 负责转录的。然而 RNA pol II 转录出 mRNA 是一个由多转录因



a、b 分别是对猪 β -actin 和 *Taf7* 基因扩增结果 (1 心脏、2 背肌、3 淋巴、4 脾脏、5 肝脏、6 肾脏、7 肺脏、8 子宫、9 睾丸、10 胃、11 小肠、12 大肠、13 卵巢、14 胸腺、15 脑、16 脂肪)

a and 4b separately show the results amplifying for β -actin and *TAF7* (1~16, represent heart, dorsal muscles, lymph nodes, spleen, liver, kidney, lung, uterus, testicle, stomach, small and large intestine, ovaries, thymus gland, brain and fat)

图 5 猪 *TAF7* 基因组织表达谱分析

Fig. 5 Tissue expression profile of the porcine *TAF7* gene

子参与的、复杂的、受到精确调节的过程。TF II D 是真核生物 RNA pol II 在转录起始识别过程中最重要成分之一，它首先识别并结合启动子的 TATA 框序列，然后吸引普通转录因子 (general transcription factors, GTFs) 结合到该位置上而构成起始转录复合物 (preinitiation complex, PIC)，最后 RNA pol II 结合到此 PIC 上开始了基因的转录启动^[15-18]。TAF7 蛋白作为 TF II D 组成成分之一，在很多基因转录起始过程中发挥重要的作用，在真核生物基因表达调控中扮演重要的角色^[19]。

人和牛 *TAF7* 基因均无内含子，均编码 349 个氨基酸；小鼠 *TAF7* 基因 5'UTR 区含有内含子，编码 341 个氨基酸。本试验第一次克隆了猪的 *TAF7* 基因，该基因 CDS (编码区) 不含内含子，只是在起始密码子上游 5'UTR 中有 3 个碱基的内含子，这一结构特征与小鼠 *TAF7* 基因一致。通过生物软件预测猪 *TAF7* 蛋白包含 349 个氨基酸残基，与人 *TAF7* 蛋白相比，氨基酸组成数目相同，而且仅有 5 个氨基酸差异，二者在系统发育树图中也聚类在一起，即这两种蛋白一级结构具备很高的相似性。蛋白质一级结构决定其高级结构，高级结构与该蛋白的功能紧密联系，系统发育树结果表明，研究猪 *TAF7* 蛋白结构和功能时，对于人 *TAF7* 蛋白功能研究结果的可借鉴性最高。同时，可以看到 *TAF7* 在物种间的保守性很高，其中有 5 个物种 (人、牛、小鼠、褐鼠和原鸡) 的变异距离小于 0.1。软件分析表明猪 *TAF7* 含有人 *TAF7* 的保守结构域，而人 *TAF7* 是转录起始过程中一个非常重要的启动子识别因子，由此推测猪 *TAF7* 在转录起始过程中发挥一定的作用。

猪 *TAF7* 基因定位于 2 号染色体 q 臂远侧端，与 SW1879 和 *IL4* 紧密连锁。已知猪 *CD14* (cluster of differentiation antigen 14, 分化抗原决定簇分子 14) 和 *GDF9* (growth differentiation factor-9, 生长分化因子-9) 基因都与 SW1879 和 *IL4* 紧密连锁^[20-21]。分析发

现，猪 *TAF7*、*IL4*、*CD14* 和 *GDF9* 基因存在连锁性，三者均定位在猪 2 号染色体上遗传距离相近的区域，这一区域与人染色体 5q2-5q3 存在保守共线性，人 *TAF7*、*IL4*、*CD14* 和 *GDF9* 基因均定位在人染色体的 5q2-5q3 区域内 (<http://genome.ucsc.edu/index.html>)，在一定程度上说明本试验结果的正确性，但尚未能获得关于 *TAF7* 与基因表达调控关系的信息。

在成年五指山猪 16 个不同组织中进行的组织表达谱分析结果表明，*TAF7* 基因可以在 16 个组织中的 14 个组织中检测得到，即该基因在大多数组织中均有表达，说明该基因具有广泛的组织表达特征，推测猪 *TAF7* 在猪生长过程中起到广泛的生物学作用^[19,22-23]。另外，该基因在睾丸中表达量较高，而在心脏和背肌中检测不到表达，推测该基因在维持雄性生理功能中发挥了某种重要作用，而在心脏和背肌其转录起始作用受到抑制，还需进一步研究探讨。

猪比其它动物在解剖学、生理学和病理学等方面与人类有更大的相似性，已成为许多人类疾病的理想动物模型^[24-26]，通过对猪 *TAF7* 基因的克隆、定位和表达谱分析，能够为研究猪精子发育机制提供依据。

4 结论

本试验首次克隆了猪 *TAF7* 基因，经序列分析后发现该基因 CDS 区域不含内含子，猪 *TAF7* 蛋白序列与人 *TAF7* 蛋白的氨基酸序列相似性较高，二者在生物系统进化过程中非常接近。猪 *TAF7* 基因定位在猪 2 号染色体 q 臂远侧端，该基因在成年五指山猪大多数组织中均检测到表达，尤其在猪睾丸中表达量较高，但在心脏和背肌中没有检测到该基因的表达。

References

- [1] Patikoglou G, Burley S K. Eukaryotic transcription factor-DNA complexes. *Annual Review of Biophysics & Biomolecular Structure*, 1997, 26: 289-325.

- [2] Hansen S K, Takada S, Jacobson R H, Lis J T, Tjian R. Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell*, 1997, 91: 71-83.
- [3] Burley S K, Roeder R G. Biochemistry and structural biology of transcription factor IID (TFIID). *Annual Review of Biochemistry*, 1996, 65: 769-799.
- [4] Pugh B F, Tjian R. Diverse transcriptional functions of the multisubunit eukaryotic TFIID complex. *Journal of Biological Chemistry*, 1992, 267: 679-682.
- [5] Rabenstein M D, Zhou S, Lis J T, Tjian R. TATA Box-Binding Protein (TBP)-Related Factor 2 (TRF2), a Third Member of the TBP Family. *Proceedings of the National Academy of Sciences of the United States of America*, 1999, 96: 4791-4796.
- [6] Hochheimer A, Tjian R. Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. *Genes & Development*, 2003, 17: 1309-1320.
- [7] Shen W C, Bhaumik S R, Causton H C, Simon I, Zhu X, Jennings E G, Wang T H, Young R A, Green M R. Systematic analysis of essential yeast TAFs in genome-wide transcription and preinitiation complex assembly. *The European Molecular Biology Organization*, 2003, 22: 3395-3402.
- [8] Gegonne A, Weissman J D, Singer D S. TAFII55 binding to TAFII250. 373. inhibits its acetyltransferase activity. *Proceedings of the National Academy of Sciences of the United States of America*, 2001, 98: 12432-12437.
- [9] Lavigne A C, Mengus G, Gangloff Y G, Wurtz J M, Davidson I. Human TAFII55 interacts with the vitamin D3 and thyroid hormone receptors and with derivatives of the retinoid X receptor that have altered transactivation properties. *Molecular and Cellular Biology*, 1999, 19: 5486-5494.
- [10] Munz C, Psichari E, Mandilis D, Lavigne A C, Spiliotaki M, Oehler T, Davidson I, Tora L, Angel P, Pintzas A. TAF7 (TAFII55) plays a role in the transcription activation by c-Jun. *Journal of Biological Chemistry*, 2003, 278: 21510-21516.
- [11] Kimmins S, Kotaja N, Davidson I, Sassone-Corsi P. Testis-specific transcription mechanisms promoting male germ-cell differentiation. *Reproduction*, 2004, 128: 5-12.
- [12] Pointud J C, Mengus G, Brancorsini S, Monaco L, Parvinen M, Sassone-Corsi P, Davidson I. The intracellular localisation of TAF7L, a paralogue of transcription factor TFIID subunit TAF7, is developmentally regulated during male germ-cell differentiation. *Journal of Cell Science*, 2003, 116: 1847-1858.
- [13] Zhou T, Chiang C M. The intronless and TATA-less human TAFII55 gene contains a functional initiator and a downstream promoter element. *Journal of Biological Chemistry*, 2001, 276: 25509-25511.
- [14] Tora L. A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription. *Genes & Development*, 2002, 16: 673-675.
- [15] Orphanides G, Lagrange T, Reinberg D. The general transcription factors of RNA polymerase II. *Genes & Development*, 1996, 10: 2657-2683.
- [16] Woychik N A, Hampsey M. The RNA polymerase II machinery: structure illuminates function. *Molecular Cell*, 2002, 108: 453-463.
- [17] Hahn S. Structure and mechanism of the RNA polymerase II transcription machinery. *Nature Structural & Molecular Biology*, 2004, 11: 394-403.
- [18] Hampsey M. Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiology and Molecular Biology Reviews*, 1998, 62: 465-503.
- [19] Leurent C, Sanders S L, Demény M A, Garbett K A, Ruhlmann C, Weil P A, Tora L, Schultz P. Mapping key functional sites within yeast TFIID. *The European Molecular Biology Organization*, 2004, 23: 719-727.
- [20] Petersen C B, Nygård A B, Fredholm M, Aasted B, Salomonsen J. Cloning, characterization and mapping of porcine CD14 reveals a high conservation of mammalian CD14 structure, expression and locus organization. *Developmental and Comparative Immunology*, 2007, 31: 729-737.
- [21] Cepica S, Procházka R, Cívánová K, Knoll A, Némecová L, Masopust M, Kubicková S, Musilová P, Rubes J. Partial molecular characterization and mapping of the GDF9 gene to porcine chromosome 2. *Animal Genetics*, 2004, 35: 245-264.
- [22] Gegonne A, Weissman J D, Zhou M S, Brady J N, Singer D S. TAF7: A possible transcription initiation check-point regulator. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, 103: 602-607.
- [23] Lavigne A C, Mengus G, May M, Dubrovskaya V, Tora L, Chambon P, Davidson I. Multiple interactions between hTAFII55 and other TFIID subunits. *Journal of Biological Chemistry*, 1996, 271: 19774-19780.
- [24] Hsu Q K. Pig model for biomedical research. *Taiwan: Pig Research Institute*, 1982: 3-10.
- [25] Bustad L K, McClellan R O. Swine in Biomedical Research. *Pacific Nothernest Laboratory Richland WA*. 1996.
- [26] Panepinto L M. Swine in biomedical research. *Laboratory Animal Science*, 1986, 36: 339-433.