

内蒙古绒山羊毛囊兴盛期皮肤 cDNA 文库 的构建及 KAP6-2 cDNA 的克隆

尹俊^{1,2}, 扈庭茂², 李金泉^{3,*}, 张春兰³, 郭志成⁴, 周欢敏¹

(1. 内蒙古农业大学 生物工程学院, 内蒙古 呼和浩特 010018; 2. 内蒙古大学 生命科学学院, 内蒙古 呼和浩特 010021; 3. 内蒙古农业大学 动物与医学学院, 内蒙古 呼和浩特 010018; 4. 内蒙古农业大学 生态环境学院, 内蒙古 呼和浩特 010018)

摘要: 用 SMART 技术构建了绒山羊 (*Capra hircus*) 毛囊兴盛期皮肤组织的 cDNA 质粒文库。Trizol Reagent (GIBCO/BRL) 分离 RNA 后, 用 Oligotex (QIAGEN) 提取 mRNA; 以锚定引物反转录合成 cDNA 第一链作为模板, 以 2 个锚定引物, 用长链 PCR 扩增全长 cDNA 双链。CHROMA SPIN-400 去除小片段后用 *Sfi* I 酶切, 连接到 *Sfi* I 消化的 pBluescript II SK (带有 *Sfi* I A 和 B 两个位点) 质粒载体中, 转化 *E. coli* 5 α , 库容量为 1.8×10^5 clones。随机挑选克隆提取质粒, 5'测定插入片段的核苷酸序列, 与 GenBank 数据库比对后发现 1 个 KAP6-2 cDNA, 序列号为 AY316158。绒山羊 KAP6-2 由 675 个核苷酸组成, 编码 96 个氨基酸, 具有高甘氨酸/酪氨酸 HGTps 与富含半胱氨酸蛋白之间的中间特征。

关键词: 内蒙古绒山羊; KAP6-2; 全长 cDNA

中图分类号: Q959.842 文献标识码: A 文章编号: 0254-5853(2004)02-0166-06

Construction of a Skin cDNA Library of Cashmere Goat and Cloning of KAP6-2 Full-length cDNA

YIN Jun^{1,2}, HU Tin-mao², LI Jin-quan^{3,*}, ZHANG Chun-lan³,
GUO Zhi-zheng⁴, ZHOU Huan-min¹

(1. College of Biotechnology, Inner Mongolia Agricultural University, Huhhot 010018, China;

2. College of life Science, Inner Mongolia University, Huhhot 010021, China;

3. College of Animal Science and Medical Science, Inner Mongolia Agricultural University, Huhhot 010018, China;

4. College of Ecosystem Environment, Inner Mongolia Agricultural University, Huhhot 010018, China)

Abstract: In order to understand the molecular mechanism of the characteristic of cashmere, we constructed a cDNA library using the SMART cDNA library construction kit (Clontech). Total RNA was isolated from the goat (*Capra hircus*) skin tissue with hair follicle anagen. Oligotex (QIAGEN) was used to isolate mRNA from total RNA. The "anchor first-strand cDNA" synthesized by reverse transcription with the SMART technique. The LD-PCR was performed using a modified oligo (dT) primer and an anchor primer as the primer set, and anchor first-strand cDNA as the template to enrich the cDNA population for full-length sequences. After digestion with *Sfi* I and size fractionation, SMART cDNA was ligated into the *Sfi* I-digested pBluescript II SK (with *Sfi* I A and B site). The ligation mixture was transformed into *E. coli* 5 α . The cDNA library contained 1.8×10^5 independent clones. Randomly select cDNA clones and sequence with the 5' prime end, a full-length KAP6-2 cDNA was found by compared with those in the NCBI database (nr) using the Blast-N programs, it showed 75.5% identity in amino acids with mouse KAP6-2 and the accession number in GenBank is AY316158.

Key words: Cashmere goat; KAP6-2; Full-length cDNA

收稿日期: 2003-10-09; 接受日期: 2004-02-02

基金项目: 国家自然科学基金地区重点项目资助 (39969002)

* 通讯作者 (Corresponding author), E-mail: lijinquanh@yahoo.com

第一作者简介: 尹俊, 内蒙古大学在职博士生; 导师扈庭茂教授。

羊绒是山羊 (*Capra hircus*) 的内层毛被, 具有轻柔、光滑, 细软等特点, 是天然纤维中的珍品, 被誉为“纤维之王”。羊绒有很好的纺织性能, 与绒纤维的蛋白组分和结构有着密切的关系。因此, 阐明绒纤维的蛋白质的组成及其基因表达对绒山羊的选育具有重要的理论和实际指导意义。目前, 动物毛发角蛋白的研究已取得了很大进展 (Powell & Rogers, 1997; Langbein et al, 2001; Fratini et al, 1993; Rogers et al, 2002), 但有关羊绒结构蛋白基因仍未见报道。为此, 我们构建了内蒙古绒山羊毛囊兴盛期皮肤 cDNA 文库, 并发现了 1 个和小鼠 KAP6-2 高度同源的序列; 为更好地了解其在绒毛生长中所起的作用, 对其结构特征进行了分析。

1 材料和方法

1.1 材料

于 10 月中旬选取阿尔巴斯雄性白绒山羊 1 只, 从体侧部剪去绒和毛, 碘酒消毒, 酒精脱碘后, 用消毒过的刀片取体侧部皮样, 立即投入液氮中保存备用。

SMARTTMcDNA 文库构建试剂盒购自 Clontech 公司, Trizol Reagent 总 RNA 分离试剂盒购自 GIBCO/BRL 公司, Oligotex mRNA MiniKit 纯化试剂盒购自 QIAGEN 公司, 菌主为 *E. coli* DH5 α 。

cDNA 合成引物: 引物均按照试剂盒要求合成, SMART (switching mechanism at 5' end of RNA transcript) III 寡核苷酸 5'-AAGCAGTGGTATCAACGCA-GAGTGGCCATTACGGCCGGG-3'; CDS III/3' PCR 引物: 5'-ATTCTAGAGGCCGAGCGGCCGACATG-d (T) 30N-1N-3'; 5' PCR 引物: 5'-AAGCAGTG-GTATCAACGCAAGT-3'; M13 (Forward-47) 通用测序引物由赛百盛公司合成。

1.2 方法

1.2.1 总 RNA 和 mRNA 的提取 山羊皮肤组织 500 mg 在液氮中研磨至粉末状, 转至匀浆管, 按 GIBCO BRL 公司的 Trizol reagent 操作说明书分离总 RNA。电泳检测总 RNA 的完整性, 紫外分光光度计测含量。

按照试剂盒说明书用 Oligotex mRNA MiniKit 从总 RNA 中分离 mRNA。紫外分光光度计测 mRNA 的含量。

1.2.2 cDNA 的合成、酶切和纯化 cDNA 第一链的合成: 按 Clontech 公司 SMARTTMcDNA 质粒文库

构建试剂盒说明书, 取 1 μ g 绒山羊皮肤 mRNA 在 5 μ L 反应体系中加入 1 μ L 10 μ mol/L CDS III/3' PCR 引物, 1 μ L 10 μ mol/L SMART III 引物, 放置于 72 $^{\circ}$ C 2 min, 冷却后加入 SuperscriptTM II 反转录酶, 10 mmol/L dNTP 混合液, 20 mmol/L DTT 1 μ L, 加入 5 \times 第一链合成 Buffer 2 μ L, 反应体积 10 μ L, 42 $^{\circ}$ C 反应 1 h 反转录合成 cDNA 第一链。双链 cDNA 的合成: 取 cDNA 第一链 2 μ L, 加入 10 μ mol/L CDS III/3' PCR 引物、5' PCR 引物各 2 μ L; 加入 50 \times Advantage 2 聚合酶混合液和 50 \times dNTP 混合液各 2 μ L, 10 \times 第二链合成 Buffer 10 μ L, 使反应体积达到 100 μ L, 用 LD PCR 扩增 cDNA 双链, 具体反应条件为: 95 $^{\circ}$ C 变性 1 min 后, 一个循环为 95 $^{\circ}$ C 变性 15 s 后 68 $^{\circ}$ C 延伸 6 min, 共进行 20 个循环。

吸取 50 μ L 扩增的 ds cDNA (2 ~ 3 μ g) 到一个 0.5 mL 灭菌的离心管中, 加入 2 μ L 蛋白酶 K (20 μ g/L), 混匀, 离心。45 $^{\circ}$ C 温浴 20 min, 离心。加入 50 μ L 去离子水, 加 100 μ L 酚: 氯仿: 异戊醇抽提, 14 000 r/min 离心 5 min。收集上层液体到另一干净的离心管中, 加入等体积的氯仿: 异戊醇混匀, 14 000 r/min 离心 5 min。收集上层液体到另一干净的离心管中, 加入 1/10 体积的 3 mol 乙酸钠, 2.5 倍体积 95% 的室温乙醇, 室温条件下立即 14 000 r/min 离心 20 min。倒掉上清, 用 100 μ L 80% 乙醇洗沉淀物, 空气干燥沉淀物约 10 min, 加 79 μ L 的去离子水溶解沉淀。在一新 0.5 mL 离心管中加入下列试剂: 79 μ L cDNA, 10 μ L 10 \times Sfi I Buffer, 10 μ L Sfi I, 1 μ L 100 \times BSA, 总体积 100 μ L, 充分混匀, 50 $^{\circ}$ C 温浴 2 h, 加 2 μ L 1% 二甲苯胺染料。于 50 $^{\circ}$ C 酶切 2 h。

1.2.3 cDNA 的分级分离和连接 准备 16 个收集管, 标记为 1 ~ 16。按 Kit 操作要求, 用 700 μ L 缓冲液洗 CHROMA SPIN-400 柱子, 待洗柱缓冲液流完, 将上面的酶切产物约 100 μ L 平稳地加到胶层中间, 直到产物全部渗到胶面下。加 100 μ L 缓冲液使染料层下降数毫米, 放置好第 1 收集管, 加 600 μ L 缓冲液, 收集滴出液体, 每管收集 35 μ L。每管取 3 μ L 用 1.1% 琼脂糖凝胶, 100 V 电泳检测 20 min。合并 cDNA 片段大于 500 bp 的第 9 ~ 13 管, -20 $^{\circ}$ C 乙醇沉淀过夜。14 000 r/min 离心 20 min。弃上清, 室温干燥 10 min。加 7 μ L 去离子水溶解沉淀后电泳检测。

在反应管中加入下列试剂: cDNA 1.0 μ L, Vec-

tor (500 ng/mL) 1.0 μ L, 10 \times Ligation Buffer 0.5 μ L, ATP (10 mmol) 0.5 μ L, T4 DNA Ligase 0.5 μ L, 灭菌三蒸水 1.5 μ L, 总体积 5.0 μ L, 16 $^{\circ}$ C 连接过夜。克隆载体为 pBluescript II SK 的改造载体, 具体是将 *Eco*R I 和 *Not* I 之间的序列改造为 *Sfi* I A 和 *Sfi* I B 接头序列, 酶切后定向插入 cDNA 片段 (*Sfi* I A \rightarrow *Sfi* I B)。

1.2.4 重组质粒的转化、重组率测定及质量鉴定

取 200 μ L 感受态宿主菌 (DH-5 α) 至灭菌的 50 mL 离心管中, 放置冰水浴融化; 加入上述 1 μ L 连接液, 轻轻混匀后置于冰上 45 min。42 $^{\circ}$ C 水浴热休克 90 s, 迅速放入冰上 2 min; 加入 2 mL LB 培养基 (不含抗生素), 37 $^{\circ}$ C 振荡培养, 转速 \leq 150 r/min, 复苏 45 min。3 000 r/min 离心 10 min 收集菌体, 用 250 μ L LB 重悬; 取 200 μ L 菌液涂布于 15 cm 培养皿上 (*Ap*^r-IPTG/X-gal LB 固体培养基), 37 $^{\circ}$ C 过夜。IPTG 和 X-gal 为上海生工生物工程技术有限公司的产品。用 T3 和 T7 测序引物, 通过 PCR 方法鉴定 cDNA 插入片段, 并估计文库的重组率。

1.3 测序及分析

随机挑取克隆, 纯化质粒 DNA, 用 M13 正向通用引物由 ABI PRISM377-96 测序仪从 5' 端测插入片段的 DNA 序列。去除载体序列和可靠性较低的尾部序列。将高质量的序列文件与 GenBank 中非冗余核苷酸数据库进行 Blastn 比较, 如果比对的期望值小于 $1e^{-10}$, 则认为两条序列来源于同一基因 (Parsons & Cooper, 1994)。合并比较结果, 形成已知功能基因文件。然后用 DNASTAR 5.01 软件中的 Clustal W 进行聚类分析和比较。

2 结果和分析

2.1 cDNA 文库的构建

高质量的总 RNA 是构建 cDNA 文库的第一步。从电泳结果看, 用 Trizol Reagent 总 RNA 分离试剂盒得到的 RNA OD₂₆₀/OD₂₈₀ 为 1.899, 大于 1.8, 说明纯度较高 (图 1)。

哺乳动物 mRNA 仅占总 RNA 的 1% ~ 5%, 因此 mRNA 的提纯会由于量的变少而非常困难。由于提取的总 RNA 有限, 分离 mRNA 后, 没有用电泳检测。本实验紫外分析 mRNA 的 OD₂₆₀/OD₂₈₀ 为 1.959, 大于 1.9, 说明纯度很高, 其纯度和含量均符合 cDNA 文库合成和建库的要求。

采用 SMART 技术, 用 0.5 μ g mRNA 反转录成

单链 cDNA, 然后 LD-PCR 扩增获得双链 cDNA, 取 5 μ L PCR 产物在 1% 的琼脂糖凝胶电泳分析 (图 2), 分子量大小为 0.1 ~ 6 kb, 长度分布均匀, 可以满足 cDNA 文库构建的要求。

cDNA 合成后经 *Sfi* I 酶切后, 在 cDNA 分级分离过程中, 收集第 9 ~ 13 管, 可保证 cDNA 文库的质量。

通过体外连接和感染宿主菌, 得到约 1 000 个克隆, 本实验所得连接产物量为 180 μ L, 菌落数约为 1 000 个, 因此库容量为 1.8×10^5 。用 T3 和 T7 测

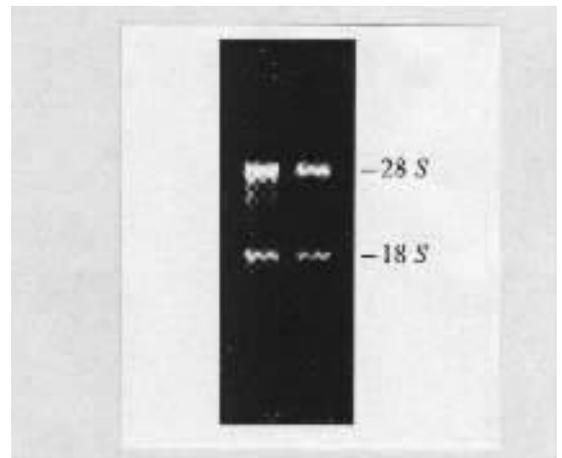


图 1 绒山羊毛囊兴盛期皮肤总 RNA 电泳图谱
Fig.1 Electrophoretic pattern of total RNA in skin of cashmere goat

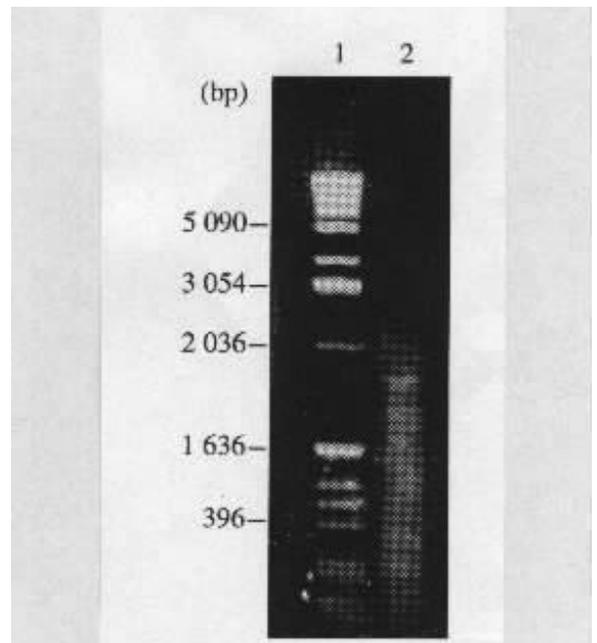


图 2 绒山羊毛囊兴盛期皮肤扩增的双链 cDNA 电泳图
Fig.2 Double strand cDNA in skin of cashmere goat
1: 1kb ladder; 2: 扩增后的双链 cDNA (ds cDNA)

序引物对挑取的 96 个单菌落进行 PCR 扩增, 很多插入片段在 0.8 kb 以上, 最长的插入片段达 2.5 kb, 其中有 12 个为空载体, 文库的重组率大于 87%。

2.2 绒山羊 KAP6-2 cDNA 的获得和比较分析

本实验共提取 861 个阳性质粒, 用 M13 正向引物从 5' 端测序, 测了 861 个, 成功 711 个, 成功率为 82.68%; ESTs 的平均长度为 543 bp, 合格的大于 100 bp 的 636 个 ESTs 的平均长度为 565 bp。

将这些序列与 GenBank 中的有关序列进行比较, 发现一个序列与小鼠 KAP6-2 同源, 确定为绒山羊 KAP-2, 其核苷酸序列及其推导的氨基酸序列见图 3。

绒山羊 KAP6-2 由 675 个核苷酸组成, $(C + G)\% = 47.96$, 拥有从 59 ~ 352 共 294 个核苷酸的开放读码框, 编码 96 个氨基酸。因为已经知道 HGTps 没有甲硫氨酸残基, 除了起始位置的甲硫氨酸 (M), 其分子量为 10 337.32 Da, 有 5 个强碱性氨基酸 (K, R), 6 个亲水性氨基酸 (A, I, L,

F, W, V), 56 个极性氨基酸 (N, C, Q, S, T, Y), 等电点为 8.096。几种 HGTp 主要氨基酸的含量分别为甘氨酸 (glycine) 27.0%, 酪氨酸 (tyrosine) 23.0%, 丝氨酸 (serine) 14.6%, 半胱氨酸 (cysteine) 15.6%, 甘氨酸和酪氨酸的含量达到 50%, 介于 I 型和 II 型 HGTp 之间。

绒山羊 KAP6-2 与小鼠 KAP6-1、KAP6-2, 人 KAP6-2, 绵羊 KAP6-1 蛋白的氨基酸序列的比对和聚类分析结果如图 4 和图 5。

由图 4 可知绒山羊 KAP6-2 和小鼠 KAP6-1、KAP6-2, 人 KAP6-2, 绵羊 KAP6-1 都含有 HGT 蛋白 N 端保守序列 MCCN (S/-) YYGN。KAP6-2 的序列 C 末端含有 H、W、V, 而其他比对序列不含这 3 种氨基酸, 从而其多肽链亲水性增强, 等电点上升呈碱性 (8.096), 此特点是绒山羊 KAP6-2 所独有的, 与 KAP-1 有明显的区别。另外, 绒山羊 KAP6-2 和小鼠 KAP6-2 的序列在 55 ~ 130 之间有明显的相似性, 小鼠 KAP6-2 从第 17 到 54 个氨基酸比山羊 KAP6-2 多 9 个序列 YG (C/S) G 的重复。

```
GACACACACTTCAGAAACATCCTCTTGCAGTCCACCTGAATCTGCTCCCCTGGACAACATGTGTTGTAACACTACTATGGCA 80
                                     M C C N Y Y G
ACTCCTGTGGCTATGGCTGTGGAACAGCTATAGCTGTGGGTTTCAGCCCCTATTATGGCTGTGGATATGGAAGTAGATAC 160
N S C G Y G C G N S Y S C G F S P Y Y G C G Y G S R Y
AGTTGTGGATATGGCTCAGGATACGGCTGTGGATACGGTTCAGGATACGGCTGTGGATATGGGACAGGATACAGCTGTGG 240
S C G Y G S G Y G C G Y G S G Y G C G Y G T G Y S C G
ATTTAGGCCCTATTATGGCTGTGGCTATGGAACCAGATATGGCTGTGGATACGGCTCAAGCTATAGCAGCCACTGGCCAG 320
F R P Y Y G C G Y G T R Y G C G Y G S S Y S S H W P
TTTGCTACAGGAGATGTTATTCTTTCTGTTAGAACATCACCGTCTTAGCCTAGTCTGCTTCTGAAATGAAGAATCTAAAG 400
V C Y R R C Y S F C *
ATAAACTGCTTTGAGGATCAGTAACCCAGATTCTACTTGCAAGAAATTACATGCCTGACAGAGACTTTTTTCTCCAA 480
CTACCCACGTTTAGAACAACGGACCTGAAGTCTGAGGTTCTATGGTGATAACATTCAAGACTCTTCCAATGTTAGCCT 560
TTACTCCCCTTAATCCTCTGAAGTCTCTTAATGAACCGGGGTAATGATCCCTAAACATTTCTATTAACATTGGAAAAA 640
CCCCCTATCCTAAATAAAAATCACTCTTCCTTCT
```

图 3 绒山羊 KAP6-2 cDNA 序列及推导的氨基酸序列

Fig. 3 Nucleotide sequence and the deduced amino acid sequence of cashmere goat KAP6-2 cDNA

下划线表示起始密码 ATG, 星号为终止密码 TAG, 框内为加尾信号 AATAAA, 基因登录号为 AY316158.

Initial codon ATG, stop cordon TAG, and A polyadenylation signal AATAAA were indicated with underline, asterisk, and bold, respectively. GenBank accession number is AY316158.

绒山羊 KAP6-2 和小鼠 KAP6-2 都有一个共同的特征,就是包括大量 (C/S)G (Y/S/C)G 的 4 个氨基酸的重复,从氨基酸序列 9 ~ 147,小鼠 KAP6-2 有 33 个 (C/S)G (Y/S/C)G 的 4 个氨基酸的重复,这些蛋白在 C 末端没有 KAP6-1 C 端的 Y 而有 C。绒山羊 KAP6-2 和小鼠 KAP6-2 的氨基酸序列的一致性最高 (75.6%),与人 KAP6-2、小鼠 KAP6-1、绵羊 KAP6-1 相似性分别为 40.3、52.6 和 48.2,均较低。从图 5 看,山羊 KAP6-2 虽然和小鼠 KAP6-2 一致最高,但并未聚为一类。

4 讨 论

毛发特异蛋白基因的克隆对研究毛发的生长和结构有重要的意义,本研究首次构建了山羊皮肤 cDNA 文库,将有助于从分子水平探讨绒毛的生长和绒毛品质相关基因的表达。对 KAP6-2 进行分析,发现它有完整的开放读码框,与其他动物 HGTps 比较有一些独特的特征:首先,绒山羊 KAP6-2 比 KAP6-1 的分子量更大,预期编码 96 个氨基酸,分子量为 10 337.32 Da;但比小鼠的 KAP6-2 分子量小。第二,绒山羊 KAP6-2 与小鼠 KAP6-2 相似,有许多 4 肽重复结构。这些重复在小鼠 HGTps II .1, II .2, II .3 和 I,以及绵羊和兔子 KAP6.1 也存在 (Rogers et al, 2002; Aoki &

Ito, 1997)。绒山羊 KAP6-2 的 4 肽重复的频率少于小鼠 KAP6-2,但多于其他的比对序列。第三,绒山羊和小鼠的 KAP6-2 相似,在 C 末端有一个半胱氨酸残基 C,而在其他 HGTps 中多为酪氨酸 (Y)。富含半胱氨酸蛋白的分子量在 10 000 ~ 30 000 Da。在氨基酸序列中包括几个重复序列,而且在羧基末端有半胱氨酸 C。这些提示绒山羊 KAP6-2 具有 HGTps 与富含半胱氨酸蛋白之间的中间特征。羊绒纤维是最好的动物纤维之一,但关于其品质特征的本质研究甚少。从毛纤维的结构蛋白分析,不同物种甚至同一物种或个体不同部位的毛纤维都有差异,这些差异都可能对毛纤维的品质产生影响。毛纤维中角蛋白关联蛋白 (KAPs) 就有几十种,其数目仍不清楚,确切的功能更是知之甚少。我们首次克隆到绒山羊高甘氨酸/酪氨酸 KAP6-2,聚类结果显示它和 KAP6 其他成员有明显差异,可能在绒纤维的生长中发挥独特的作用。当然,它在绒山羊绒和毛生长中的表达特点和功能还有待于进一步深入研究。

致谢:感谢阿尔巴斯种羊场刘少卿厂长、张永斌、苗雄、扬二虎、白小梅、李润平等技术员在采样工作中提供的大力协助!

参考文献:

- Aoki N, Ito M. 1997. Isolation and characterization of mouse high-glycine/tyrosine proteins [J]. *J. Biol. Chem.*, **272** (48): 30512 - 30518.
- Fratini A, Powell BC, Rogers GE. 1993. Sequence, expression and evolutionary conservation of a gene encoding a glycine/tyrosine-rich keratin-associated protein of hair [J]. *J. Biol. Chem.*, **268**: 4511 - 4518.
- Langbein L, Rogers MA, Winter H, Ehmann C, Praetzel S, Schweizer J. 2001. The catalog of human hair keratins [J]. *J. Biol. Chem.*, **276** (37): 35123 - 35132.
- Parsons YM, Cooper DW. 1994. Evidence of linkage between high-glycine-tyrosine keratin gene loci and wool fibre diameter in a Merino half-sib family [J]. *Anim. Genet.*, **25** (2): 105 - 108.
- Powell BC, Rogers GE. 1997. The formation and structure of human hair [M]. Basel: Birkhäuser Verlag. 59 - 148.
- Rogers MA, Langbein L, Winter H, Ehmann C, Praetzel S, Schweizer J. 2002. Characterization of a first domain of human high glycine-tyrosine and high sulfur keratin-associated protein (KAP) genes on chromosome 21q22.1 [J]. *J. Biol. Chem.*, **277**: 48993 - 49002.