

五指山小型猪 *APEX1* 基因 cDNA 的克隆及生物信息学分析李治深¹, 杜丽¹, 刘涛¹, 申明霞^{1,2}, 王凤阳^{1*}, 成鹰¹, 陈品林¹, 林杰材¹, 满初日嘎¹, 祁超²

(1. 海南大学农学院, 海口市动物基因工程重点实验室, 海南海口 570228; 2. 华中师范大学生命科学学院, 湖北武汉 430079)

摘要 [目的] 对五指山小型猪脱嘌呤嘧啶核酸内切酶 cDNA 进行克隆和生物信息学分析。[方法] 以构建的五指山小型猪外周血白细胞 cDNA 文库为材料, 采用菌落 PCR 方法, 克隆得到 *APEX1* 基因全长 cDNA, 并运用生物信息学软件对该基因核苷酸序列及其编码蛋白的理化性质及二级结构等进行分析 and 预测。[结果] 生物信息学分析表明, 该 cDNA 全长 1 392 bp, 5' 非翻译区长 132 bp, 3' 非翻译区长 303 bp, 含有一个 957 bp 的完整开放阅读框, 编码 318 个氨基酸。该蛋白的分子量为 35 kD, 等电点为 8.05。比对分析表明, 五指山小型猪 *APEX1* 基因的核苷酸序列及其氨基酸序列与人、小鼠、黑猩猩等哺乳动物具有较高的相似性。[结论] 成功克隆了五指山小型猪 *APEX1* 基因 cDNA, 为进一步研究 *APEX1* 在动物体内的作用机制奠定了基础。

关键词 五指山小型猪; 脱嘌呤嘧啶核酸内切酶; cDNA 文库; 克隆; 序列分析

中图分类号 S828 文献标识码 A 文章编号 0517-6611(2009)30-14621-04

cDNA Cloning and Bioinformatics Analysis of *APEX1* Gene from Wuzhishan Miniature Pig

LI Zhi-shen et al (College of Agriculture, Hainan University, Animal Genetic Engineering Key Lab. of Haikou, Haikou, Hainan 570228)

Abstract [Objective] The study aimed to clone full-length cDNA of *APEX1* gene from Wuzhishan miniature pig and bioinformatics analysis be carried out of this gene. [Method] Taking constructed cDNA library of the peripheral blood leucocytes from Wuzhishan miniature pig as material, the full-length cDNA of *APEX1* gene was cloned by using colony PCR method. Nucleotide sequence of *APEX1* gene, physicochemical properties and secondary structure of its encoding protein were analyzed and predicted by using bioinformatics software. [Result] The bioinformatics analysis results showed that the full-length cDNA sequence was 1 392 bp, containing an open reading frame of 957 bp, flanked by 132 bp 5'UTR and 303 bp 3'UTR, encoded 318 amino acids. Relative molecular weight was 35 kD, pI was 8.05. Comparison analysis indicated that the nucleotide sequence and amino acid sequence of *APEX1* gene were highly homologous to some mammals reported, such as human, mouse, chimpanzee and other mammalian. [Conclusion] The cDNA of *APEX1* gene from Wuzhishan miniature pig was successfully cloned, it lays a foundation for further researching the action mechanism of *APEX1* in the body of animal.

Key words Wuzhishan miniature pig; *APEX1*; cDNA library; Cloning; Bioinformatics analysis

五指山小型猪是我国特有的小型猪。具有性成熟早、遗传稳定、代谢率低、适应性和抗逆性强等特点, 是理想的动物模型和异种器官移植供体。

脱嘌呤嘧啶核酸内切酶 (Apurinic/aprimidinic endonuclease, *APEX1*) 是一种多功能的 DNA 碱基切除修复途径 (Base excision repair, BER) 的限速酶, 是细胞 DNA 放射性损伤和烷化剂致伤的重要修复因子。细胞内外的物理、化学或生物刺激, 都能使细胞 DNA 染色体脱嘌呤、嘧啶而被损伤, *APEX1* 则具有修复这种损伤的功能^[1]。*APEX1* 可以调节多种细胞反应的关键转录因子, 如 p53、HIF2 α 、NF2KB 和 AP21 等^[2]。其中, HIF2 α 通过 VEGF 依赖途径和 VEGF 非依赖途径影响血管生成, p53、NF2kb、AP21 也可通过增加、促进肿瘤血管生成因子的表达和降低、抑制肿瘤血管生成因子的表达来调节肿瘤血管的生成^[3]。

笔者获得了五指山小型猪 *APEX1* 基因的 cDNA 序列, 并分析其核苷酸序列和氨基酸序列, 为进一步研究 *APEX1* 在动物细胞的新陈代谢、机体防御等生理活动中的作用和功能提供了理论依据和基本的研究材料。

1 材料与方

1.1 材料

1.1.1 试验材料。五指山小型猪外周血白细胞 cDNA 文库

由该实验室构建。

1.1.2 试剂。 $2 \times Taq$ PCR MasterMix 购自天根生化科技(北京)有限公司。PCR 鉴定引物: M13 正向引物: 5'-GTA-AAACGAC GGCCAG-3' 和 M13 反向引物: 5'-CA GGAAA-CAGCATATGAC -3' 由上海生工生物工程有限公司合成。其他常用试剂由上海生工生物工程有限公司提供。

1.2 方法

1.2.1 菌落 PCR。 将文库菌液稀释, 涂布于含有氯霉素抗性的 LB 琼脂糖培养基上, 用灭菌的牙签挑取单菌落, 以 M13 引物进行菌落 PCR。PCR 反应程序: 94 $^{\circ}$ C 变性 3 min, 94 $^{\circ}$ C 变性 1 min, 65 $^{\circ}$ C 退火 1 min, 72 $^{\circ}$ C 延伸 30 s, 35 个循环; 72 $^{\circ}$ C 延伸 5 min, 4 $^{\circ}$ C 保存 24 h, 测序鉴定。

1.2.2 生物信息学分析^[4-5]。采用 DNAMAN、BIOEDIT 等软件及 ExPasy 站点的各种分析工具对 *APEX1* 核苷酸序列和推演的氨基酸序列进行比较和分析; 并预测蛋白质的理化性质、亲水性轮廓、跨膜信息、二级结构和三级结构等。

2 结果与分析

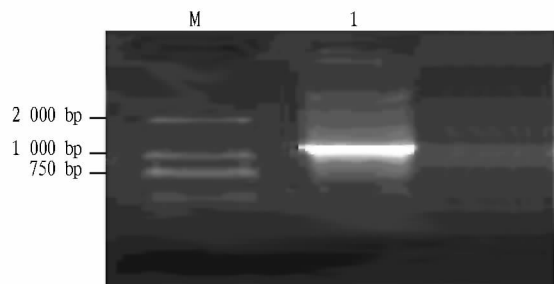
2.1 五指山小型猪 *APEX1* 基因 cDNA 的克隆 用 M13 引物进行菌落 PCR, 电泳结果见图 1。由图 1 可知, 该片段长约 1 300 bp 左右 (含载体序列 445 bp), 经测序并将全长 cDNA 序列与 GenBank 数据库进行核苷酸同源性比较, 结果表明, 该基因与人 *APEX1* 的 CDS 相似性为 89.66%, 确定五指山小型猪 *APEX1* 基因为目的基因片段。

2.2 五指山小型猪 *APEX1* 基因 cDNA 及氨基酸序列特征分析 五指山小型猪 *APEX1* 基因核苷酸序列及其推导的氨基酸序列见图 2。由图 2 可知, 该 cDNA 序列全长 1 392 bp, 拥有一个 957 bp (133 ~ 1 089 nt) 的开放读码框, 编码 318 个

基金项目 国家自然科学基金项目 (30560021); 海南省自然科学基金项目 (30509); 海南省教育厅高校科研项目 (Hj200804) 资助。

作者简介 李治深 (1950 -), 男, 海南海口人, 高级兽医师, 从事动物功能基因组学研究。* 通讯作者, 教授, 博士生导师, E-mail: lywang68@yahoo.com。

收稿日期 2009-06-26



注: M; DL-2000 Marker; 1; 菌落 PCR 产物。

Note: M. DL-2000 marker; 1. Colony PCR product of *APEX1* gene.

图 1 五指山小型猪 *APEX1* 基因 cDNA 菌落 PCR 结果

Fig. 1 Result of colony PCR of *APEX1* gene cDNA from Wuzhishan miniature pig

氨基酸,是一个完整性较好的全长 cDNA。通过 http://www.expsy.org/tools/pi_tool.html 网站, Compute pI/Mw tool 分析, 该蛋白理论等电点 (PI) 及分子量 (MW) 分别为 8.04 和 35 412.34 D。

通过 DNAMAN 软件分析五指山小型猪 *APEX1* 蛋白的氨基酸组成特性,发现该编码蛋白由 318 个氨基酸残基组成,其中,碱性氨基酸 48 个,酸性氨基酸 39 个,极性氨基酸 93 个。

2.3 五指山小型猪 *APEX1* 基因 cDNA 的生物信息学分析

2.3.1 五指山小型猪与其他动物 *APEX1* 基因的同源性比较。运用 BLAST,通过对数据库的检索,对该基因 CDS 区及其推导的氨基酸序列与 GenBank 上公布的人、小家鼠、褐家

```

1      GGG GGC TCT AGA GAG GAG TCT TTC CCC AGC CCT AGC AGG TTT CAC GAC TTT TTT GCG TCT
61     GTA GGC AAC GCG GTA GAA ACA CTG CTT CGG TGG GCG ACG TGG AAC TGC TAC AGA AGG GCG
121    TTT ATT ACA GTG ATG CCG AAA CGT GGG AAA AAA GCC GCG GTG GCA GAA GAT GGS GAT GAG
1      M P K R G K K A A V A E D G D E
181    CCC AAG TCT GAG CCA GAA GCC AAG AAG AGT AAG GCG GGA GCA AAG AAA ACC GAA AAA GAG
17     P K S E P E A K K S K A G A K K T E K E
241    GCA GCA GGT GAG GGG GCA GTC CTG TAT GAG GAC CCC CCA GAT CAG AAA ACC TCA CCC AGT
37     A A G E G A V L Y E D P P D Q K T S P S
301    GGC AAA TCA GCT ACA CTC AAA ATC TGC TCC TGG AAT GTG GAT GGG CTT CGA GCT TGG ATT
57     G K S A T L K I C S W N V D G L R A W I
361    AAG AAA AAA GGT TTA GAT TGG GTA AAG GAA GAA GCC CCA GAT ATC CTG TGC CTC CAA GAG
77     K K K G L D W V K E E A P D I L C L Q E
421    ACC AAA TGT TCA GAG AAC AAA CTG CCA GCT GAA CTT CAA GAG CTG TCT GGA CTA CCC CAT
97     T K C S E N K L P A E L Q E L S G L P H
481    CAG TAC TGG TCA GCT CCT TCT GAC AAG GAA GGG TAC AGT GGT GTG GGC CTG CTC TCC CGC
117    Q Y W S A P S D K E G Y S G V G L L S R
541    CAG TGC CCA CTC AAA GTC TCC TAT GGC ATT GGT GAG GAG GAA CAT GAT CAG GAA GGC CGA
137    Q C P L K V S Y G I G E E E H D Q E E G R
601    GTG ATT GTG GCT GAA TTT GAT GCA TTT GTG CTG GTG ACA GCC TAT GTA CCC AAT GCA GGC
157    V I V A E F D A F V L V T A Y V P N A G
661    CGA GGT CTG GTA CGC CTG GAG TAC CGC CAG CGC TGG GAT GAA GCC TTT CGC AAA TTC CTG
177    R G L V R L E Y R Q R W D E A F R K F L
721    AAG GGT CTG GCC TCC CAC AAG CCC CTT GTG CTC TGT GGG GAC CTT AAT GTG GCT CAT GAA
197    K G L A S H K P L V L C G D L N V A H E
781    GAA ATT GAC CTT CGC AAC CCA AAG GGA AAC AAA AAG AAT GCT GGC TTC ACT CCA CAG GAA
217    E I D L R N P K G N K K N A G F T P Q E
841    CGT CAA GGC TTT GGG GAA TTG CTG CAG GCT GTG CCA CTG GCT GAC AGT TTC CGG CAC CTC
237    R Q G F G E L L Q A V P L A D S F R H L
901    TAC CCC AAC ACG GCG TAT GCC TAC ACC TTT TGG ACC TAC ATG ATG AAT GCG CGA TCT AAA
257    Y P N T A Y A Y T F W T Y M M N A R S K
961    AAT GTT GGA TGG CGG CTT GAT TAT TTT TTG TTG TCT CAC TCT CTG TTG CTT GCA CTG TGT
277    N V G W R L D Y F L L S H S L L P A L C
1 021  GAC AGC AAG ATC CGT TCC AAG GCT TTG GGT AGT GAC CAC TGT CCC ATT ACT CTA TAC CTA
297    D S K I R S K A L G S D H C P I T L Y L
1081  GCA CTG TGA CCT TTC CCC AAA TCA GTT TGA GCC TGG GAT ATA AGC CAA TCA AAC TAG CTT
317    A L *
1 141  TAA ACT TCT TAC GCC CAA ACT TTA AAA ACT GCT CTC TAG AGA AAT TTG CAC TGT ATT TTC
1 201  CTT TTC AAC TGC AGC TTT TAA CCA GGC TTC TAG TAA CAG ATG TAG GTT TCC TTT AAG CCC
1 261  AAG GAT TTT TGT GTG TGT TGG GAT GTC TTT TTT TTT TTA ATA TTA AAA ACT ACT GAT GAG
1 321  TTC CTT TGA AAT GCC CAT GTG AAA ATA AAA AGC CAT AGT TTC GAA AAA AAA AAA AAA AAA
1 381  AAA AAA AAA AAA
    
```

图 2 五指山小型猪 *APEX1* 基因 cDNA 序列及推导的氨基酸序列

Fig. 2 *APEX1* gene cDNA and deduced amino acid sequence cloned from Wuzhishan miniature pig

鼠、黑猩猩的 *APEX1* 基因 CDS 序列的同源性及其氨基酸序列同源性进行了比对(图 3)。CDS 序列比对分析表明,五指山小型猪 *APEX1* 与褐家鼠 (NM_024148)、小家鼠 (NM_009687)、人 (NM_080648)、黑猩猩 (NM_001081485) 同源性分

别为 85.79%、86.10%、89.66%、89.97%。氨基酸比对分析表明,五指山小型猪 *APEX1* 的氨基酸序列与小家鼠同源性高达 95.60%,其次为黑猩猩 94.97%、人 94.65%、褐家鼠 94.65%。

小家鼠	MPKRGKK. AAADDGEEPKESEPETKKSCKGAAKKTEKEAAGEGVPVLYEDPPDQKTSPPSGKSATLKICSWNVDGLRAWIKKKG	79
褐家鼠	-----r.-----e-----a-----	79
黑猩猩	-----g-v-e-d-lrt-a---t---nd-----a-----p-----	80
人	-----g-v-e-d-lrt-a---t---nd-----a-----p-----	80
五指山猪	-----a-v-e-d-----a---ag-----a-----	80
小家鼠	LDWVKEEAPDILCLQETKCSENKLPALQELPGLTHQYWSAPSDKEGYSVGLLSRQCPLKVSYGIGEEHDQEGRVIVA	159
褐家鼠	-----	159
黑猩猩	-----s-----	160
人	-----s-----d-----	160
五指山猪	-----s p-----	160
小家鼠	EFESFVLVTAYVPNAGRGLVRLEYRQRWDEAFRKFCLKDLASRKPLVLCGDLNVAHEEIDLNRNPKGNKKNAGFTPQERQGF	239
褐家鼠	-----i-----	239
黑猩猩	---d-----g-----	240
人	---d-----g-----	240
五指山猪	---da-----g---h-----	240
小家鼠	GELLQAVPLADSRHLYPNTAYAYTFWYMMNARSKNVNGWRLDYFLLSHSLLPALCDSKIRSKALGSDHCPITLYLAL	317
褐家鼠	---m-----	317
黑猩猩	-----p-----	318
人	-----p-----	318
五指山猪	-----	318

图 3 五指山小型猪和其他哺乳动物的 APEX1 编码的氨基酸序列比对结果

Fig.3 Alignment of the induced amino acid sequence encoded by APEX1 of Wuzhishan miniature pig and other mammal

2.3.2 五指山小型猪 APEX1 蛋白跨膜区、疏水性及信号肽分析。分别利用 <http://www.cbs.dtu.dk/services/TMHMM/> 网站的 TMHMM Server v. 2.0 程序和 BIOEDIT 软件进行蛋白序列跨膜区分析和疏水性分析(图 4)。跨膜区分析表明,该序列无跨膜区,表明不是跨膜蛋白。疏水性分析表明,氨

基酸序列中亲水性和疏水性区域间隔存在,小部分为中性。分值在 +1.2 ~ +1.8 的 1 处,在 160 ~ 170,富含疏水性氨基酸。<http://www.cbs.dtu.dk/services/SignalP/> 网站分析表明,此序列可能不含信号肽。

2.3.3 五指山小型猪 APEX1 蛋白二级结构的预测。进入

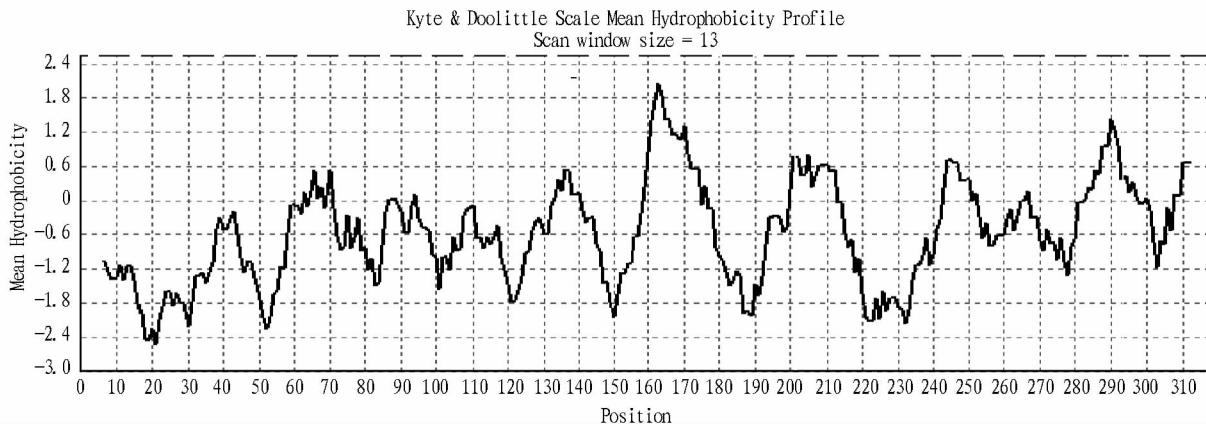


图 4 五指山小型猪 APEX1 蛋白亲/疏水性分析结果

Fig.4 The hydrophilic/hydrophobic analysis results of Wuzhishan miniature pig's APEX1 protein

http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hnn.html 网址,运用 PBIL LYON-GERLAND 信息库对蛋白质序列进行二级结构预测(Secondary structure prediction),主要运用 Hopfield 神经网络(HNN)进行预测。预测结果表明,组成五指山小型猪 APEX1 蛋白的 318 个氨基酸中,87 个氨基酸可能形成 α 螺旋,59 个氨基酸可能形成 β 折叠片,172 个氨基酸可能形成无规卷曲(图 5)。

2.3.4 五指山小型猪 APEX1 蛋白质三级结构的预测。该试验利用 SWISS-MODEL 同源建模服务器 <http://swissmodel.expasy.org> 进行蛋白质三级结构的预测,结果见图 6。

3 结论讨论

APEX1 除具有核酸内切酶活性,参与 DNA 损伤修复外,还具有氧化还原功能,通过维持细胞内多种转录因子的激活还原态,参与多种关键的细胞反应,如氧化应激、转录因子调节、细胞周期控制与凋亡。研究发现,在肺癌、乳腺癌、前列腺癌、宫颈癌等多种肿瘤中发现 APEX1 过表达,并与肿瘤细胞的恶性转化、浸润和转移密切相关,另外,DNA 修复机制的紊乱被认为是癌症演进的发病机制^[2,6-8]。

有学者提出假说认为,APEX1 在细胞内的定位不同与 DNA 修复程序相关的特定功能有关,表达定位的改变以及表

