# 猪泛素 C 端水解酶 L1 基因的克隆与序列分析

田兴国,李加琪,陈瑶生,高 萍,梅盈洁,凌 飞,罗永发 (华南农业大学 动物科学学院, 广东 广州 510642)

摘要:从人泛素 C 端水解酶  $L1(UCH_L1)$ 基因出发,在 dbEST 数据库中同源搜索,找到 1 条与人  $UCH_L1$  基因编码 氨基酸同源性较高且在香猪背最长肌中表达的 EST(BM194679). 通过电子克隆和进一步 RT - PCR 试验验证,获得 猪 UCH\_L1 基因全长 cDNA 序列,其全长 1 105 bp,开放阅读框(ORF)位于 60~728 bp,编码 223 个氨基酸. 同源性 分析结果表明,与人、鼠 UCH\_L1 基因 cDNA 编码区(CDS)同源性分别为 91.2% 和 86.5%,蛋白质序列同源性均为 96.6%.对该基因编码蛋白的结构和功能预测显示含有2个典型的疏水性区域,不含有信号肽,存在1个UCH\_L1 (pfam01088)保守结构域和多个磷酸化位点,属 UCH\_L1 蛋白家族,故将该基因命名为猪泛素 C 端水解酶 L1 基因 (登录号 AY495532).

关键词:猪; 泛素 C 端水解酶  $L1(UCH_L1)$  基因; 电子延伸; 基因序列分析

中图分类号:S813.1

文献标识码:A

文章编号:1001-411X(2007)01-0100-05

## Cloning and Sequence Analysis of Ubiquitin C-Terminal Hydrolase L1Gene of Pig

TIAN Xing-guo, LI Jia-qi, CHEN Yao-sheng, Gao Ping, MEI Ying-jie, LING Fei, LUO Yong-fa (College of Animal Sciences, South China Agric. Univ., Guangzhou 510642, China)

Abstract: The human ubiquitin C-terminal hydrolase L1 (UCH\_L1) gene cDNA was used to search the dbEST database for homologous porcine expressed sequence tags (ESTs). A matching EST (BM194679) previous found to express Xiang pig longissimus dorsi muscle was identified. Through in silico cloning and RT-PCR with total RNA from Landrace pig longissimus dorsi muscle as the template, a 1 105 bp cDNA fragment was amplified and subsequently sequenced. Sequence analysis revealed that this 1 105 bp fragment contained a 669 bp open reading frame and shared the highest homogenecitis at 91.2% and 86.5% with human and mouse homologues, respectively. The deduced peptide of 223 amino acid residues showed 96.6% homogenety to the orthologous of human and mouse, and contained a classical UCH\_L1 (pfam01088) domain, 2 typical transmembrane domains and 3 biological-meaning sites, but with no signal peptide. Based on these characteristics this gene was nominated as pig UCH\_L1 gene (GenBank acession number AY495532).

Key words: pig; ubiquitin C-terminal hydrolase L1 (UCH L1) gene; in silico cloning; gene sequence analysis

hydrolase L1,缩写为 UCH\_L1)是一种广泛分布的蛋

泛素 C 端水解酶 L1 (ubiquitin carbroxy-terminal 白质,它属于泛素 C 端水解酶家族,参与泛素依赖的 蛋白水解通路,具有水解泛素分子和其他蛋白之间

收稿日期:2006-05-15

作者简介:田兴国(1976--),男,硕士,现工作单位:华南农业大学科学技术处;通讯作者:李加琪(1965--),男,教授,博士, E-mail:jqli@ scau. edu. cn

肽键的功能,因此 UCH\_L1 对于提供单体泛素分子以供标记目标蛋白使其被水解具有重要作用<sup>[1]</sup>. UCH\_L1 蛋白在人脑组织中含量比较丰富(占 2%),在骨骼肌中也含有 UCH\_L1 蛋白,它参与泛素 - 蛋白酶体蛋白质水解途径并在骨骼肌蛋白代谢过程中发挥重要的作用. 研究发现了 UCH\_L1 基因的错义突变(Ile93Met)可使该巯基蛋白酶的催化活力下降,从而造成蛋白分解和蛋白积累的错误,从而造成骨骼肌萎缩和病变<sup>[2]</sup>. 笔者从人 UCH\_L1 基因出发,通过电子克隆并结合 RT - PCR 技术从猪骨骼肌中扩增得到猪 UCH\_L1 基因,为以后研究猪瘦肉蛋白形成的分子基础提供重要的依据.

### 1 材料与方法

#### 1.1 材料

组织样本来自广东省韶关市畜牧局种猪场长白猪背最长肌组织,各个体均为健康无病的肥育猪;菌株大肠杆菌 DH-5α 由华南农业大学动物遗传育种与繁殖实验室保存;质粒 pUCm-T 由上海生物工程有限公司提供; TRIZOL Reagent RNA 提取试剂盒购自GibcoBRL公司; DEPC, RNasin, DNase I 购自 Promega公司; ExTaq DNA 聚合酶, dNTPs, TaKaRa RNA PCR (AMV) Ver2.1 试剂盒,核糖核酸酶抑制剂(Rbonuclease inhibitor), DNA maker DL 2000,购自 TaKaRa 公司; TA 克隆试剂盒购自上海生物工程有限公司.

#### 1.2 方法

- 1.2.1 总 RNA 的抽提和检测 取长白猪背最长肌组织样,TRIZOL 法提取总 RNA,20 g/L 琼脂糖电泳验证 RNA 的完整性,并用 Eppendorf 核酸蛋白浓度测定仪检测 RNA 的浓度.
- 1.2.2 EST 电子克隆和分析 以人 *UCH\_L1* 基因 mRNA 序列为模板,利用 BLASTN 程序,上 NCBI (http://www.ncbi.nlm.nih.gov)网站检索 dbEST 数据库,找到 1 条同源性较高的香猪背最长肌 EST (BM194679),以这条 EST 为种子序列对 UniGene 数据库检索.将重叠序列大于 100 bp 且相似性在 90%以上的 ESTs 用 DNAstar 软件拼接、延伸,电子延伸后的序列进行开放阅读框(ORF)识别<sup>[3]72-77</sup>.
- 1.2.3 ORF 的 RT PCR 验证和基因序列分析 在 ORF 起始密码子和终止密码子两侧设计 1 对引物——LF 和 LR. 引物设计采用 Genetool 软件,引物序 列: LF (5'-AAACCACATCCAAAGGCCTTAAAGT-CACAG 3') 和 LR (5'-GCGAAGATGCAGCTCAAAC-CGATG -3');引物由上海生物工程有限公司合成. 以

提取的总 RNA 为模板,用 TaKaRa A-MLV 酶反转录 cDNA,并用  $\beta$ -actin 加以验证. 以反转录 cDNA 为模 板合成第 2 条链. 2.5  $\mu$ L 10 × ExTaq PCR buffer, 1  $\mu$ L 50 pmol/L 上、下游混合引物,2  $\mu$ L 25 mmol/L MgCl<sub>2</sub>,2  $\mu$ L 2.5 mmol/L dNTPs,反转录混合物 1.5  $\mu$ L,ExTaq 酶 0.75 U,加 双 蒸 水 至 终 体 积 为 25  $\mu$ L. PCR 循环参数:94  $\Omega$  变性 60 s,然后 94  $\Omega$  变性 30 s,57  $\Omega$  退火 30 s,72  $\Omega$  延伸 90 s,32 个循环,最后 72  $\Omega$  延伸 8 min. 得到的 PCR 产物经柱式 胶回收试剂盒进行纯化、回收,送上海博亚生物技术 有限公司进行 PCR 产物克隆测序. 测序后的序列在 NCBI 上用 ORF-finder 软件进行分析并用 BLASTN 程序进行基因序列及 BLASTP 蛋白质序列的同源性 检索  $\Omega$  是  $\Omega$  是

1.2.4 蛋白质结构、功能分析 登录 http://www.expasy.org/cgi-bin/protscale.pl 网站应用位于 Ex-PASy 的 ProtScale 程序进行蛋白质的疏水性图谱分析, 登录 http://genome.cbs.dtu.dk/services/TM-HMM-2.0,进行蛋白质序列的跨膜区分析, 联网至http://www.genome.cbs.dtu.dk/services/SignalP/对猪 UCH\_L1 蛋白序列进行信号肽分析,并登录 http://www.expasy.org/prosite 的 PROSITE 数据库检索蛋白质的功能位点;蛋白质的功能预测利用 BLASTP软件进行相应分析[3]120-136.

### 2 结果与分析

#### 2.1 猪 UCH\_L1 基因电子克隆结果

以人 UCH\_L1 基因编码区序列为模板,在美国 NCBI 网站上利用 BLASTN 程序检索 EST 数据库,搜寻到 1 条与人 UCH\_L1 基因编码氨基酸高度同源的香猪背最长肌 EST(BM194679). 以 BM194679 序列为种子序列,利用 UniGene 数据库进行电子延伸得到 1 105 bp 片断. 在靠近 3′和 5′两端的地方设计特异引物进行 RT – PCR 扩增以验证电子克隆后的序列,预计扩增片断长度为 1 000 bp.

#### 2.2 RT-PCR 验证

利用设计的 1 对特异引物(LF,LR),采用 RT - PCR 技术,从抽提的长白猪背最长肌组织总 RNA 反转录 cDNA 中扩增出 1 000 bp 左右的片断(图 1),与预期的扩增片断长度相符,且特异性与重复性较好.

#### 2.3 阳性克隆的鉴定

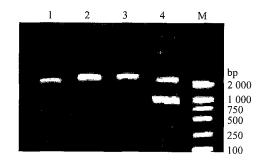
将目的基因克隆到 pUCm - T 载体,转化大肠杆菌 DH -  $5\alpha$ ,挑选阳性的克隆,PstI 酶切进行鉴定(图 2),结果与预期的一致,初步确认已获得猪  $UCH_LI$ 

基因 cDNA 序列.



M:DNA marker DL2000;1: 据 UCH\_LI 基内的 RT-PCR产物 M:DNA marker DL2000;1: UCH\_LI gene RT-PCR product of pig

图 1 猪 UCH-L1 基因的 RT - PCR 产物 Fig. 1 RT-PCR product of UCH L1 gene



M:DNA marker DL2000; 1、2、3:阴性; 4:阳性
M:DNA marker DL2000; 1,2,3: negative; 4: postive
图 2 重组质粒的酶切鉴定

Fig. 2 Identification of recombinant plasmid by enzyme digestion product of pig

#### 2.4 目的基因序列测定和分析

将上述筛选的阳性克隆送上海博亚生物技术有限公司测序,测序结果在美国 NCBI 上进行 ORF 和同源性分析. 猪 *UCH\_L1* 基因全长 cDNA 1 105 bp (图3),包含完整的 ORF.5′端上游含有 59 bp 的非

编码区,ORF 位于 60~728 bp 之间,长度为 669 bp, 编码 223 个氨基酸;60~62 bp处是起始密码子 ATG, 729~731 bp处是终止密码子 TAA,731 bp后为 3′端 非编码序列.在 ORF 起始码上游同一相位有终止码 TAG,ORF 后有一个加尾信号和 PolyA 尾. 该序列提 交 GenBank 收录(登录号为 AY495532).

利用 BLASTN 和 BLASTP 程序进行同源性分析表明,该序列与人、鼠 UCH\_L1 核苷酸序列有较高同源性,其中与人和鼠 UCH\_L1 核苷酸序列同源性分别为 79.7% 和 76.3%,编码区序列同源性为91.2%、86.5%,人、鼠 UCH\_L1 编码区长度均为672 bp,与人、鼠 UCH\_L1 编码氨基酸序列同源性均为96.6%(图4),证明该基因编码的蛋白质序列比较保守.

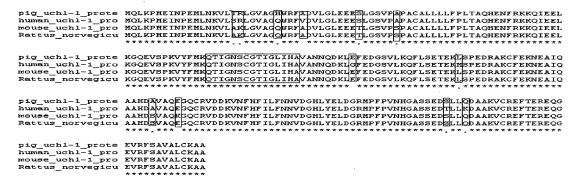
#### 2.5 蛋白质基本性质

应用位于 ExPASy 的 ProtScale 程序(http://www.expasy.org/cgi-bin/protscale.pl) [3]120 来计算蛋白质的疏水性图谱(图 4),分析结果显示在 30~60位之间和 79~105位之间含有 2 个典型的疏水性区域.联网至 http://genome.cbs.dtu.dk/services/TM-HMM-2.0 [3]121,进行蛋白质序列的跨膜区分析(图 5)结果表明,在第 39~57位之间存在 1 个由外向内的跨膜区,在 80~100位之间含有 1 个由内向外的跨膜区,与该蛋白质的疏水性区域分析结果基本一致.联网至 http://www.genome.cbs.dtu.dk/services/SignalP/[3]124 对猪 UCH\_L1 蛋白序列进行信号肽分析,分析结果显示该蛋白序列不含有信号肽,故断定该蛋白是一个非分泌性蛋白.

 $\tt GCGGCCGCTAGATGTTTTTCGTCTTCTTTGGGTTGCGGYYSCTCTCGGCTTTT$ 

加下划线的为加尾信号; "\*"标记的为终止密码子; 阴影示引物序列
Polyadenylation signal is underlined; Stop codon is indicated by "\*"; Primer sequence is shadowed
图 3 UCH\_L1 基因的 cDNA 序列和推测的蛋白质序列

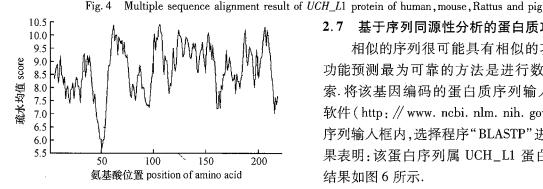
Fig. 3 cDNA and predicted protein sequence of UCH\_L1 gene



蓝色方框为 UCH\_L1 蛋白结构域,"\*"表示氨基酸一致,红色框和"."表示氨基酸不一致

Blue box contains the UCH\_L1 domain, Identical amino acid is indicated by " \* "; Similar but not identical amino acid is indicated by red box and "."

### 图 4 人、家鼠、挪威鼠和猪 UCH\_L1 基因蛋白质序列多重比对结果



### 图 5 猪 UCH\_L1 基因编码的蛋白疏水性分析 Fig. 5 Hydrophobic profile of protein encoded by pig UCH\_L1

2.6 功能位点分析结果

采用 http://www.expasy.org/prosite[3]134 的 PROSITE 数据库检索发现,该蛋白质序列中存在 1 个有生物学意义的模式(pattern) UCH\_L1(泛素 C 末端水解酶 L1 蛋白家族) 以及 3 个有生物学意 义的位点(sites)(共9个),其中包括3个蛋白激 酶 C 磷酸化位点、2 个 N-端豆蔻酰化位点和 4 个 酪蛋白磷酸化位点.

#### 2.7 基于序列同源性分析的蛋白质功能预测结果

相似的序列很可能具有相似的功能,蛋白质的 功能预测最为可靠的方法是进行数据库相似性检 索. 将该基因编码的蛋白质序列输入 NCBL/BLAST 软件(http://www.ncbi.nlm.nih.gov/blast/)[3]52的 序列输入框内,选择程序"BLASTP"进行相应分析结 果表明:该蛋白序列属 UCH\_L1 蛋白家族. BLASTP 结果如图 6 所示.

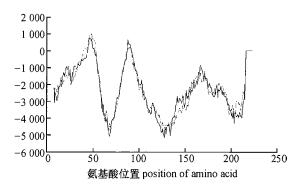


图 6 猪 UCH\_L1 基因编码的蛋白跨膜区分析 Fig. 6 Transmembrane profile of protein encoded by pig UCH\_L1



snl|CDD|25685 pfam01088, Peptidase\_C12, Ubiquitin carboxyl-terminal hydrolase, family 1. CD-Length = 212 residues, 100.0% aligned Score = 211 bits (539), Expect = 5e-56

QLKPMBINPEMLNKVLTRLGVAGHWRFADVLGLEERSLGSVPAPAGALLLLFPLTAQHENRWLPLBSNPGVFTBLLHKLGVKG-WQVEDVYSLDDBLLEMVPRPVHALIPLFPWTEKVE-62 59 FRKKQIEBLK--GQEVSPKVYFMKQTIGNSCGTIGLIHAVANNQDKLEFEDGSVLKQFLS FRKEEPKGIKVQDQDVSDSVFFAKQTIPNACGTQALLHALANNEE-IDLDLGSTLKRFLE ETEKLSPEDRAKCFEKNEAIQAAH--DAVAQEGQC---RVDDKVNFHFILFNNVDGHLYEFTKDLSPEBRGKALENDEEIRAAHNSFARAQPGQTEAPEADEDVNYHFIAYVPINGTLYE

LDGRMPFPVNHGASSEDSLLQDA-AKVCREFTERE LDGRKEGPINHGPTSDDTLWNDAVRPVIKERMERY

红色为保守氨基酸;蓝色为半保守氨基酸;黑色为不保守氨基酸 Red letter is 100%-aligned, blue letter is 50%-aligned, black letter is unaligned amino acid 图 7 基于序列同源性分析的蛋白质功能预测结果

Fig. 7 Prediction result of protein encoded by pig UCH\_L1 by blast

### 3 讨论

哺乳动物体内蛋白的降解代谢有溶酶体内和溶 酶体外2个系统. 泛素 - 蛋白酶体(ubiquitin proteasome)途径属于溶酶体外蛋白降解系统.溶酶体外蛋 白降解系统是近20年来发现的高效细胞内蛋白降 解系统,它是由泛素(ubiquitin)以及一系列的相关酶 组成[4]. 泛素是泛素 - 蛋白酶体蛋白降解途径中重 要的组成部分,其存在于所有有核细胞中[5]. 相关的 酶包括泛素活化酶(ubiquitin-activating enzyme, E1)、 泛素耦连酶(ubiquitin-conjugating enzymes, E2s)、泛 素 - 蛋白连接酶 (ubiquitin-protein ligatin enzymes, E3s)、泛素末端水解酶(ubiquitin-end hydrolysis)或 同工肽酶(isopeptidase)和蛋白酶体(proteasome)[6]. 泛素 C 末端水解酶是泛素末端水解酶的一种,参与 体内蛋白质水解代谢过程. 在生理条件下,骨骼肌与 其他真核细胞一样,蛋白质降解途径较复杂,它们包 括 ATP - 泛素途径、溶酶体途径和依赖 Ca2+ 途径 等[7]. 其中溶酶体途径主要通过组织蛋白酶 B、D、H 和 L 等起作用. 依赖 Ca2+ 途径主要通过 Ca2+ 激活中 性蛋白酶 Calpain I 或 II 起作用. 而依赖 ATP - 泛素 途径则是真核细胞内主要的降解系统. 在 ATP 存在 时,泛素与细胞内的底物蛋白共价结合,并将需降解 的蛋白提呈给 26S 蛋白酶复合体降解. Lowell 等[8] 1994年报道骨骼肌中的肌原纤维主要是通过依赖 ATP - 泛素系统降解,并且与溶酶体途径的作用有交 叉重叠. 最近有报道表明,溶酶体中存在泛素与蛋白 的结合物,且泛素与溶酶体膜的形成密切相关[9].提 示溶酶体途径是与依赖 ATP - 泛素系统相互作用 的. 由于溶菌酶膜破裂时组织蛋白酶 B、D、H 和 L 等 释放对肉的嫩化起作用. 所以推测泛素 - 蛋白酶体 蛋白水解途径对猪瘦肉蛋白质合成和肉的嫩度均有. 一定的影响, 而 UCH\_L1 是这个途径的重要酶, 对 UCH\_LI 的研究将有助于了解猪瘦肉蛋白质合成和 肉的嫩化机制分子基础.

本研究利用网上序列信息结合分子生物学实验技术快速获得了猪 UCH\_L1 基因,序列分析发现其为该基因全长 cDNA,与人、鼠该基因的同源性较高.编码区长度保守,且编码蛋白同源性高达 96.6%,表明该基因在哺乳动物中较为保守.对其编码蛋白质进行结构功能分析发现含有 2 个典型的跨膜区,存在多个磷酸化位点,并含有 UCH\_L1 蛋白家族的保守域.含有跨膜区提示它可能作为膜受体起作用,也可能是定位于膜的锚定蛋白或者离子通道蛋白等.包含 UCH\_1 蛋白家族的保守域,提示属于 UCH\_L1

家族成员,与该蛋白质家族其他成员具有相似的生物学功能.磷酸化是蛋白质合成后广泛存在的一种化学修饰,是调控酶活性的重要步骤.在许多情况下,磷酸化以后酶活性大大提高,从而对蛋白质的功能调节起重要作用,猪 UCH\_L1 蛋白具有多个磷酸化位点,可见它在蛋白降解过程中受到精细调节,同时具有重要作用.本研究应用现有公认比较权威的核酸、蛋白质序列分析数据库及软件对得到的猪新基因 UCH\_L1 进行一系列结构功能预测,与同类数据库分析结果一致,证实分析的准确性.

本研究报道了猪 UCH\_L1 基因全长 cDNA 序列,有利于研究该基因的分子调控机制以及为运用 RNA 干扰、定点突变等方法研究基因功能提供了可能,将有助于研究猪肌肉蛋白合成和肌肉嫩化的分子机制.

#### 参考文献:

- [1] HERSHKO A, CIECHANOVER A. The ubiquitin system for protein degradation [J]. Annu Rev Biochem, 1992, 61: 761-807.
- [2] LEROY E, BOYER R, AUBURGER G, et al. The ubiquitin pathway in Parkinson's disease [J]. Nature, 1998, 395:451-452.
- [3] 张成岗,贺福初. 生物信息学方法与实践[M]. 北京: 科技出版社,2002.
- [4] KAJIMOTO, HASHIMOTO T, SHIRAI Y, et al. cDNA cloning and tissue distribution of a rat ubiquitin carboxylterminal hydrolase PGP9. 5[J]. J Biochem, 1992, 112:28-32
- [5] WILKINSON K D, LEE K M, DESHPANDE S, et al. The neuron-specific protein PGP 9.5 is a ubiquitin carboxylterminal hydrolase [J]. Science, 1989, 246:670-673.
- [6] ARNOLD J, DAWSON S, FERGUSSON J, et al. Ubiquitin and its role in neurodegeneration [J]. Prog Brain Res, 1998, 117:23-34.
- [7] KERTTELHUT I C, PEPALO M T, MIGLIORINI R H, et al. Regulation of different proteolytic pathway sinskeletal muscle in fasting and diabetes mellitus [J]. Bragillian J Med Bio Res, 1994, 27:981-993.
- [8] LOWELL B B, RUDERMAN N B, GOODMAN M N. Regulation of myofibrillar protein degradation in rat skeletal muscle during brief and prolonged starvation[J]. Metabolism, 1986, 35:1121-1127.
- [9] SCHQARTH A L, BRANDT R A, GEUGE H, et al. Stress induce dalterations in auto phagic pathway relationship to ubiquitin system [J]. Am J Physiol, 1992, 262: 1031-1038.

【责任编辑 柴 焰】