

鹅 LPL 基因外显子4、5、6 克隆及序列分析

于建兴^{1,2} 陈宏权^{1,3} 邱华玲¹ 殷宗俊¹

(1. 安徽农业大学 动物科技学院, 合肥 230036; 2. 嘉兴学院 生化学院, 浙江 嘉兴 314001;
3. 湖南农业大学 动物科技学院, 长沙 410128)

摘要 为揭示鹅脂蛋白脂酶(LPL)脂质和肝素结合区域的特性,对鹅 LPL 基因外显子 4~6 进行了克隆和序列分析。以皖西白鹅血清总 DNA 为模板,设计引物对鹅 LPL 基因的第 4~6 外显子区域进行扩增和测序,分析鹅与其他物种之间该区域的 DNA 序列同源性、密码子非随机应用以及相应氨基酸残基区域的亲水性。结果表明,鹅 LPL 基因外显子 4~6 的片段大小分别为 112、234 和 243 bp,其序列与鸡 LPL 基因相应序列的同源性达到 91.9%,与人、羊、牛、猪等哺乳动物的同源性为 74%~77%;鹅与鸡密码子非随机应用的相似系数为 0.839,与猪和鼠的相似系数为 0.580~0.650;LPL 中由外显子 4~6 编码的氨基酸残基有 4 个区域具有较大的表面概率,分别位于残基 145~148、残基 187~190、残基 255~257 和残基 293~299 区域,表面概率分别 1.62~4.29、2.06~3.01、2.75~4.46 和 3.03~6.01,这些区域伴随有较强的亲水性。

关键词 皖西白鹅; 脂蛋白脂酶; 外显子; 基因克隆; 序列分析

中图分类号 Q 785; S 835 文章编号 1007-4333(2007)05-0010-06 文献标识码 A

Cloning and analyzing exon 4, 5 and 6 of lipoprotein lipase gene in goose

Yu Jianxing^{1,2}, Chen Hongquan^{1,3}, Qiu Hualing¹, Yin Zongjun¹

(1. College of Animal Science and Technology, Anhui Agricultural University, Hefei 230036, China;
2. Biochemical College of Jiaxing University, Jiaxing 314001, China;
3. Animal Science and Technology College, Hunan Agricultural University, Changsha 410128, China)

Abstract Exon 4 to intron 6 sequence of lipoprotein lipase (LPL) gene was cloned to determine genome characteristics in goose. The total DNA from goose serum was extracted. Two pairs of primers, P and Q, were designed to amplify the LPL gene exon 4 to intron 6 which were then cloned and sequenced. The results showed that the LPL gene exon 4 to intron 6 were 3 404 bp. Blasting the sequence homology among chicken, human, pig, mouse, cattle and cat showed that the DNA homology between goose and chicken was 91.9% which was the highest between goose and other species. The homology between goose and mammal (human, sheep, cattle and pig) was 74 to 77%. Also the hydrophilicity, average charge, surface probability and condon usage of amino acids coded by LPL gene exon 4 to exon 6 in goose were analysed.

Key words goose; lipoprotein lipase; exon; gene clone; sequence analysis

脂蛋白脂酶(lipoprotein lipase, LPL)是由脂肪细胞、心肌细胞、骨骼肌细胞、乳腺细胞及巨噬细胞等实质细胞分泌的一种酰基甘油水解酶,催化水解乳糜微粒和极低密度脂蛋白中的甘油三酯,转化成非酯化脂肪酸和 2-酰基甘油提供给组织^[1]。脂肪

组织中的 LPL 活性升高将有助于机体脂质的贮存,而骨骼肌中的 LPL 活性升高则有助于机体产热^[2]。LPL 在脂肪组织中的活性及在肌肉组织中的相对活性决定了脂肪的储存或供能,于是 LPL 基因已经作为动物脂肪代谢和沉积以及畜禽肉质研究的重要

收稿日期: 2006-05-29 修回日期: 2007-06-18

基金项目: 安徽省“十五”人才基金资助项目(人才-4);安徽省教育厅自然科学基金资助项目(2006 KJ 169B)

作者简介: 于建兴,硕士研究生;陈宏权,教授,通讯作者,主要从事动物分子数量遗传学研究, E-mail: chqchq@ahau.edu.cn, chenhq62@126.com

基因^[3-5]。目前灵长类、啮齿类、鱼类以及家畜中部分动物的 *LPL* 基因被克隆和测序,不同动物的 *LPL* 基因均由 10 个外显子和 9 个内含子组成,其中第 1 外显子包含 5 端非翻译区和 *LPL* 的信号肽,第 4 外显子编码脂质结合区域,第 6 外显子编码肝素结合区域^[6-7],外显子 2~5 主要编码酶底物结合区和催化活性中心区域,而外显子 6~9 序列编码区域参与将酶锚定在血管内皮表面^[6]。禽类 *LPL* 基因的研究比其他动物起步较晚,除鸡^[8]的 *LPL* 基因被克隆和测序以外,其他禽类 *LPL* 基因的研究尚未见报道。为此,本研究通过克隆鹅 *LPL* 基因的外显子 4~6,并进行序列分析,为进一步研究鹅脂类代谢和脂肪沉积遗传规律提供分子生物学基础。

1 材料与方法

1.1 材料

本研究所用试验鹅来自于皖西白鹅与四川白鹅杂交试验的亲本群体,按照皖西白鹅的饲养标准饲养;随机抽取 10 只鹅(雌雄各半)的血液提取 DNA 样品。扩增用 Taq 酶购自上海生物工程公司,转化载体 pMD18-T 购自大连宝生物公司。

1.2 引物设计与合成

在 GenBank 中获得鸡 *LPL* 基因全长序列,利用外显子的高同源性,用 Primer Premier 5.0 软件设计 2 对(P 和 Q)扩增目的片段的引物:

P 上游引物:

ATT GTC ACT TGC TGG GGT ACA GTC T

下游引物:

GAC ATC CAC GAA GTC AGC GTC AT

Q 上游引物:

CTG GTC CCA CCT TTG AGT ATG C

下游引物:

TCT CGT CTA GAG TGC CAT ACA GAG A

引物 P 用于扩增目的片段的前段,覆盖外显子 4 和内含子 4;引物 Q 用于扩增目的片段的后段,覆盖外显子 5、内含子 5、外显子 6 和内含子 6。P 和 Q 有交叉克隆区域,引物均由上海生物工程公司合成。

1.3 鹅基因组 DNA 的提取

采用鹅翅下静脉采血,血样加入 ACD 抗凝剂于 -20℃ 保存。以酚-氯仿法^[9]提取血清基因组总 DNA。DNA 经紫外分光光度计检测,其 D_{260}/D_{280} 为 1.751。

1.4 PCR 体系与反应条件

以提取的鹅基因组总 DNA 为模板,采用 50 μ L 的 PCR 反应体系:10 \times PCRBuffer 4 μ L, Mg^{++} (25 mmol/L) 4 μ L, dNTP (2.5 mmol/L) 4 μ L, DNA 提取物 1.1 μ L, 上下游引物各 0.8 μ L (125 ng/ μ L), TaqDNA 聚合酶 (10 U/ μ L) 0.5 μ L。反应条件如下:94℃ 预变性 10 min, 94℃ 45 s, 57℃ 45 s, 72℃ 2 min 30 s, 经 35 个循环后,在 72℃ 延伸 10 min。取 6 μ L 反应液在 1.2% (质量分数) 琼脂糖凝胶上电泳,分析 PCR 产物大小。

1.5 目的片段测序

用 V-gene 凝胶回收试剂盒回收目的片段;参照氯化钙法制备 DH5⁺ 感受态细胞^[10];分别取上述纯化后的 PCR 产物 4 μ L, 与 pMD18-T 载体相连,然后转化入感受态细胞中,经蓝白斑筛选获得可能插入目的 DNA 的白色菌落,培养后菌液用原设计引物进行验证。送菌液到上海生物工程公司进行测序。

1.6 序列分析

本研究分析用软件为 Primer Premier 5.0 和 DNASTar 5.01。

1) 同源性分析。将鹅的 *LPL* 氨基酸序列与 GenBank 中(括弧中附收录号)的人 (*Homo sapiens*, human, BC011353)、狒狒 (*Papio anubis*, olive baboon, U18091)、猪 (*Sus scrofa*, pig, NP-999451)、牛 (*Bos taurus*, cattle, XP-616349)、羊 (*Ovis aries*, sheep, NP-001009394)、狗 (*Canis familiaris*, dog, XP-859720)、貂 (*Mustela vison*, American mink, CAA11411)、猫 (*Felis catus*, cat, U42725)、鼠 (*Mus*, house mouse, NP-032535) 和鸡 (*Gallus gallus*, chicken, P11602) 10 个物种进行比较,分析外显子 4~6 的 DNA 序列之间的同源性。本研究采用 DNASTar 软件的 MegAlign 过程获得不同序列之间的相似率和分歧率,比较不同序列之间的同源性。

2) 密码子使用频率。利用鹅 *LPL* 基因外显子 4~6 所对应的氨基酸序列,分析其氨基酸的组成和密码子的应用频率。密码子使用频率 = $\frac{mk}{n}$, 其中 m 为某种氨基酸同义密码子的种类数; n 为该种氨基酸在蛋白质中的使用次数; k 为该种氨基酸某种密码子的使用次数。当某种氨基酸的同义密码子随机使用,则每种密码子的使用频率等于 1,用密码子使用频率与 1 之间的差异大小来衡量密码子使用频率的偏差。

3) 参数分析。

氨基酸的亲水性:根据 Kyte 等^[11]的氨基酸亲水性得分绘制鹅 LPL 氨基酸亲水性分布图;

电荷密度分布:采用 DNASTar 软件绘制鹅 LPL 氨基酸残基的电荷密度分布;

氨基酸特定区域位于蛋白质表面的概率:假如内部没有有意义的缺失和插入,对某个氨基酸序列而言,位置 N 处于表面的概率是由位置 N - 2 和 N + 3 所决定的。氨基酸序列位置 N 位于蛋白质表面的概率(S)可以表示为

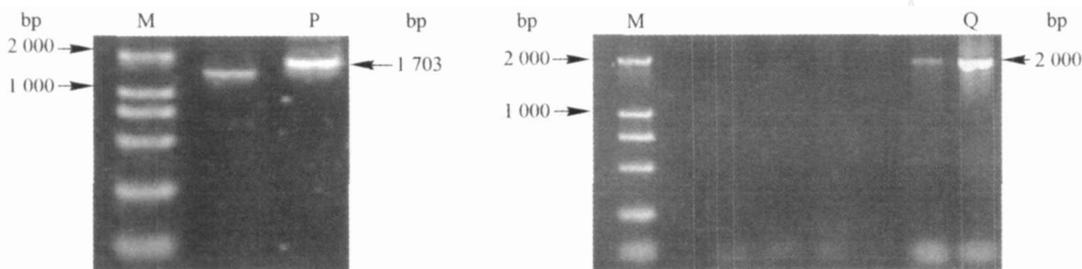
$$S_n = 0.37^{-6} \prod_{i=1}^6 p_{n+4-i}$$

其中 p_x 是位置 x 位于蛋白质表面的概率,当某一六聚序列的 S_n 大于 1,则氨基酸区域位于蛋白质表面的概率就越大^[12]。

2 结果与分析

2.1 LPL 基因外显子 4~6 序列的分析测定

根据 PCR 扩增和凝胶电泳结果,引物 P 扩增片段为 1 703 bp,引物 Q 扩增片段为 2 000 bp(图 1)。



M 为 Marker DL2000;P 和 Q 为引物 P 和 Q 的 PCR 产物带

图 1 目的片段 PCR 产物凝胶电泳图

Fig. 1 Electrophoresis of PCR production by oligo P and Q

对所扩增的目的片段经纯化、转化、蓝白斑筛选菌落过程,菌液测序(结果登录在 GenBank 上,收录

号为 DQ105796 和 DQ105797),其中外显子 4~6 的序列如下:

```

gagcagttca actacctct caacaatgct cacttgctgg ggtacagtct gggtgctcac
gctgctggga ttgctgggag cctgaccaag aagaaggtga acagaattac tgggctggat
cctgctggtc ccaccttga gtgtgctgat gccctcacac gcctctcccc ggatgacgct
gacttcgtgg atgtcctaca cacctacacc cgggctctc cggatcgcag catcggcatt
cagaagcccc taggacacat cgatatttac cctaacgggg gaggctcca gccgggctgc
aacttgggag aaactctacg cctgattgct gaaaaaggct ttgcagatgt ggatcagctg
gtgaaatgct ctcacgaacg atccatccac ctcttcateg actccctct ctacgaagaa
aagcccagca tggcctaccg ctgcaacacg aaggaggcct ttgagaagg gctctgcctc
agctgccgca agaaccgctg caacaacttg ggttacaagg tcaacagagt gagaacaaag
agaaacacca aaatgtacct gaagaccgt gctcagatgc cctacaaag
    
```

2.2 不同物种的同源性

根据序列测定结果,比照鸡的外显子 4~6 序列,鹅 LPL 基因外显子 4~6 的长度分别为 112、

234 和 243 bp。鹅 LPL 基因 3 个外显子与鸡的同源性最高,达到 91.9%,与其他物种之间的同源性只有 74%~77%(表 1)。

表 1 鹅与其他 10 个物种 LPL 基因外显子 4~6 的同源性

Table 1 Homology comparison of exon 4 to 6 in LPL genes between goose and 10 other species

项 目	鸡	人	狒狒	鼠	羊	猫	貂	狗	牛	猪
相似率	91.9	77.1	76.8	76.6	76.3	75.6	75.4	74.9	74.6	74.4
分歧率	8.7	27.7	28.2	28.5	29.0	30.3	30.5	31.1	31.7	31.8

2.3 外显子 4~6 编码的氨基酸序列分析

酸如下(表 2)。

鹅外显子 4~6 编码 *LPL* 蛋白质的 196 个氨基

EQFN YPLNNV HLL GYSL GAH AA GIA GSL TK KKVNRITGLD PAGPTFECAD AL TRLSPDDA
DFVDVL HTYT R GSPDRSIGI Q KPV GHIDIY PNGGGFQPGC NL GEALRL IA E KGFADVDQL
V KCSHERSIH L FIDSLLYEE KPSMA YRCNT KEAFEKGLCL SCRKNRCNNL GYKVNVRVTK
RNTKM YL KTR AQMPYK

表 2 鹅 *LPL* 中由外显子 4~6 编码的氨基酸组成

Table 2 Composition of amino acids coded by exon 4 to 6 of *LPL* gene

氨基酸	代号	数量	比例/ %	氨基酸	代号	数量	比例/ %
Ala	A	14	7.14	Pro	P	10	5.10
Cys	C	7	3.57	Gln	Q	5	2.55
Asp	D	11	5.61	Arg	R	13	6.63
Glu	E	9	4.59	Ser	S	10	5.10
Phe	F	7	3.57	Thr	T	10	5.10
Gly	G	17	8.67	Trp	W	0	0
His	H	6	3.06	Ile	I	9	4.59
Val	V	9	4.59	Tyr	Y	9	4.59
Lys	K	15	7.65	极性氨基酸		53	27.04
Leu	L	20	10.20	疏水性氨基酸		59	30.10
Met	M	3	1.53	带电荷氨基酸		70	35.71
Asn	N	12	6.12	酸性氨基酸		20	10.20
				碱性氨基酸		28	14.29

分析密码子使用频率(表 3)可以发现,不同的氨基酸采用不同的密码子使用策略,Ala 偏向使用 GCU,Arg 使用 CGC,His 使用 CAC,Leu 使用 CUC +

CU G,Ser 使用 AGC,Val 使用 GUG,Thr 使用 ACC 等,显示出密码子使用的非随机性。比较鹅与其他 10 个物种该区域密码子的使用策略,结果见表 4。

表 3 鹅 *LPL* 中由外显子 4~6 编码的氨基酸密码子使用频率

Table 3 Codon usage frequencies of exon4 to 6 of *LPL* genes in goose

| 氨基酸 密码子使用频率 |
|---------------|---------------|---------------|---------------|---------------|
| Ala GCA 0.286 | Gly GGA 0.760 | Pro CCA 0.000 | Val GUA 0.444 | Thr ACA 0.800 |
| GCC 0.857 | GGC 1.176 | CCC 1.600 | GUC 1.333 | ACC 2.400 |
| GCG 0.000 | GGG 1.412 | CCG 1.200 | GUG 2.222 | ACG 0.400 |
| GCU 2.857 | GGU 0.706 | CCU 1.200 | GUU 0.000 | ACU 0.400 |
| Arg AGA 1.846 | Leu CUA 0.600 | Ser AGC 2.400 | Asn AAC 1.833 | Gln CAA 0.000 |
| AGG 0.000 | CUC 2.400 | AGU 0.600 | AAU 0.167 | CAG 2.000 |
| CGA 0.462 | CUG 2.100 | UCA 0.000 | Asp GAC 0.545 | Glu GAA 1.111 |
| CGC 3.231 | CUU 0.000 | UCC 1.800 | GAU 1.455 | GAG 0.889 |
| CGG 0.000 | UUA 0.000 | UCG 0.000 | Cys UGC 1.714 | AUA 0.000 |
| CGU 0.462 | UUG 0.900 | UCU 1.200 | UGU 0.286 | Ile AUC 1.333 |
| His CAC 2.000 | Lys AAA 0.553 | Phe UUC 1.143 | Tyr UAC 2.000 | AUU 1.667 |
| CAU 0.000 | AAG 1.467 | UUU 0.857 | UAU 0.000 | |

表4 鹅与其他10个物种 *LPL* 基因外显子4~6编码的氨基酸密码子使用频率的相似性
Table 4 Correlation of codon usage of exon 4 to 6 in *LPL* genes between goose and other species

物种	鸡	猪	鼠	羊	人	狒狒	牛	貂	猫	狗
鹅	0.839	0.649	0.579	0.506	0.487	0.475	0.441	0.359	0.273	0.144

由表4可知,鹅的密码子使用策略与其他物种有明显的区别,与食肉类动物(狗、猫和貂)的密码子使用频率的相似系数较低(小于0.360),与人、牛、羊和狒狒的相似系数为0.440~0.510,与猪和鼠的相似系数为0.580~0.650,与鸡的相似系数达0.839,表明鸡和鹅在该区域的密码子使用策略相似。

2.4 *LPL* 中外显子4~6编码区域的亲水性、电荷密度及表面概率

外显子4~6编码 *LPL* 的氨基酸残基在117~312区域,表面概率有4个明显峰值区域(图2):残基145~148、残基187~190、残基255~257和残基293~299,表面概率分别为1.62~4.29、2.06~3.01、2.75~4.46和3.03~6.01,电荷密度分别为

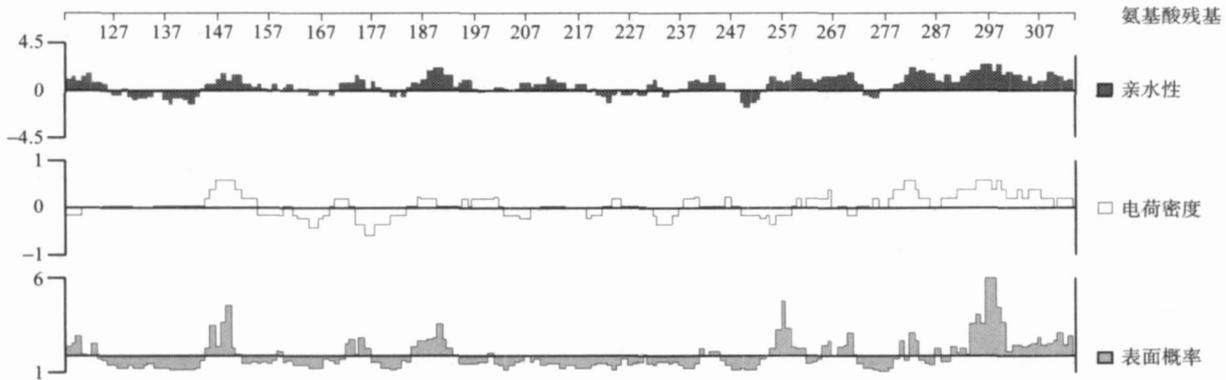


图2 鹅 *LPL* 的残基117~312区域的亲水性、电荷密度与表面概率

Fig. 2 Hydrophilicity, average charge, and surface probability of amino acid residue 117-312 of *LPL* in goose

0.4~0.6、0.0~0.2、-0.2和0.4~0.6,这些区域伴随着较强的亲水性,表明是 *LPL* 的脂质和肝素的结合区域。

3 讨论

1) *LPL* 基因是一个长度约为35 kb的单拷贝基因,由10个外显子和9个内含子组成,其中第4外显子编码脂质结合区域,第6外显子编码肝素结合区域^[6-7]。本研究分析的外显子4~6就是编码鹅 *LPL* 肝素结合区和磷酸基团结合区域。序列分析表明,鹅与鸡 *LPL* 基因外显子4~6的DNA序列长度一致,其同源性高达91.9%,与哺乳动物的同源性介于74%~77%之间,略高于鸟类与哺乳动物cDNA全长序列的同源性(61%)^[13-14],表明外显子4~6涉及到 *LPL* 的重要结合区域,不同物种之间具有更高的保守性。

2) 本研究分析的外显子4~6编码 *LPL* 氨基酸残基117~312,其中由外显子5编码的 *LPL* 氨基酸

残基157-160和187-196区域具有很小的表面概率,属于疏水链区域,这证实了该区域是构成 *LPL* 环疏水链的报道^[15]。由外显子6编码的 *LPL* 氨基酸残基292~299拥有较高的表面概率和亲水性,与该区域是构成 *LPL* 肝素结合区域的报道相符^[16]。

3) 证据表明,同义密码子或“简并”密码子在各种生物的许多基因中是以非随机方式应用的,体现出一种应用规则而不是一个例外^[17]。本研究分析显示,鹅外显子4~6编码的氨基酸密码子也存在非随机应用的规则,并且编码策略与鸡的相似性高达0.839,与其他哺乳动物之间的相似性偏低,反映出鹅与其他物种 *LPL* 基因的根本差异与密码子使用有着密切的关联性。

参 考 文 献

- [1] Wang C S, Hartsuck J, McConathy W J. Structure and functional properties of lipoprotein lipase[J]. *Biochimica et Biophysica Acta*, 1992, 1123(1):1-17

- [2] James R M, Scott A I, Dipak P R. Lipoprotein lipase: structure, function, regulation, and role in disease [J]. *Journal of Molecular Medicine*, 2002, 80(12):753-769
- [3] 潘爱奎,皮敬松,梁振华,等. 6个品种鸡 LPL 基因的 PCR-RFLP 分析 [J]. *中国家禽学报*, 2004, 8(1):152-154
- [4] 张保军,杨公社,孙超,等. 猪 LPL 基因 PCR-RFLP 遗传变异的研究 [J]. *西北农林科技大学学报(自然科学版)*, 2005, 33(3):5-10
- [5] Iakunina N I, Shestakova M V, Voron'ko O E, et al. Polymorphic gene markers of lipid metabolism are associated with diabetic nephropathy in patients with type 1 diabetes mellitus[J]. *Genetika*, 2005, 41(7):931-937
- [6] Saera-Vila A, Calduch-Giner J A, Gómez-Requeni P, et al. Molecular characterization of gilthead sea bream (*Sparus aurata*) lipoprotein lipase. Transcriptional regulation by season and nutritional condition in skeletal muscle and fat storage tissues[J]. *Comparative Biochemistry and Physiology-Part B: Biochemistry & Molecular Biology*, 2005, 142(2):224-232
- [7] Oka K, Tkalcevic G T, Nakano T, et al. Structure and polymorphic map of human lipoprotein lipase gene [J]. *Biochimica et Biophysica Acta*, 1990, 1049(1):21-26
- [8] Cooper D A, Lu S C, Viswanath R, et al. The structure and complete nucleotide sequence of the avian lipoprotein lipase gene [J]. *Biochimica et Biophysica Acta*, 1992, 1129(2):166-171
- [9] 卢圣栋. 现代分子生物学实验技术[M]. (第2版). 北京:中国协和医科大学出版社, 1999
- [10] 萨姆布鲁克,拉塞尔. 分子克隆实验指南[M]. (第3版). 北京:科学出版社, 2002
- [11] Kyte J, Doolittle R F. A simple method for displaying the hydropathic character of a protein [J]. *Journal of Molecular Biology*, 1982, 157(1):105-132
- [12] Emimi E A, Hughes J, Perlow D, et al. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide [J]. *Journal of Virology*, 1985, 55(3):836-839
- [13] Enerback S, Gimble J M. Lipoprotein lipase gene expression: physiological regulators at the transcriptional and posttranscriptional level [J]. *Biochimica et Biophysica Acta*, 1993, 1169(2):107-125
- [14] Raisonnier A, Etienne J, Arnault F, et al. Comparison of the cDNA and amino acid sequences of lipoprotein lipase in eight species [J]. *Comparative Biochemistry and Physiology-Part B: Biochemistry & Molecular Biology*, 1995, 111(3):385-398
- [15] Klaus A D, Helen L D, Silvia S F. Human Hepatic and Lipoprotein Lipase: The Loop Covering the Catalytic Site Mediates Lipase Substrate Specificity [J]. *Journal of Biological Chemistry*, 1995, 270(43):25396-25401
- [16] Lookene A, Nielsen M S, Glemann J, et al. Contribution of the carboxy-terminal domain of lipoprotein lipase to interaction with heparin and lipoproteins [J]. *Biochemical and Biophysical Research Communications*, 2000, 271(1):15-21
- [17] Kimura M. The neutral theory of molecular evolution [M]. Cambridge:Cambridge University Press, 1983