

李丽莎,李祥龙,毛艳朋. 山羊 *PMEL* 基因生物信息学分析[J]. 江苏农业科学,2016,44(8):47-50.
doi:10.15889/j.issn.1002-1302.2016.08.011

山羊 *PMEL* 基因生物信息学分析

李丽莎¹, 李祥龙^{1,2}, 毛艳朋¹

(1. 河北农业大学,河北保定 071000;2. 河北科技师范学院,河北秦皇岛 066004)

摘要:基于生物基因组学数据库,对山羊 *PMEL* 基因进行生物信息学分析,为研究该基因的功能提供依据。根据 GenBank 中山羊 *PMEL* 基因编码蛋白序列 (GenBank 登录号:XP_005680442.1),利用 ExPASy 数据库中的 ProtParam、ProtScale 程序分别预测理化性质、疏水性;利用 CBS Prediction Servers 数据库中的 SignalP、TMHMM、Protfun 程序分别预测信号肽、跨膜区及其功能;利用 Predict Protein、PSORT II Prediction、Motif Scan 软件分别预测二级结构、亚细胞定位、功能位点,并对不同物种 *PMEL* 基因构建系统发育树。结果表明:山羊 *PMEL* 基因共编码 685 个氨基酸,属可溶性不稳定蛋白,无信号肽,在接近肽链的 C-端存在 1 个跨膜区;二级结构以无规卷曲为主,该蛋白主要在内质网中发挥运输、结合、信号转导作用;不同物种 *PMEL* 基因的系统进化情况与生物进化过程中的动物学分类相符。山羊 *PMEL* 基因编码蛋白在生物进化中具有较强保守性,该基因结构和功能的分析为山羊 *PMEL* 基因遗传特性的进一步研究奠定了基础。

关键词:PMEL;山羊;生物信息

中图分类号: Q785;S826.9⁺2

文献标志码: A

文章编号: 1002-1302(2016)08-0047-04

前黑素小体蛋白 (premelanosome protein, PMEL) 别称 gp100、ME20、PMEL17、silver^[1],由 *Silver* 基因编码,只在皮肤黑色素细胞、眼色素层黑色素细胞、视网膜色素上皮细胞中进行表达^[2]。前黑素小体蛋白是一种黑色素瘤特异性糖蛋白,对黑素小体的发育具有重要作用,通过形成蛋白水解的纤维状矩阵从而使黑色素沉积^[3]。人类 *PMEL* 基因编码蛋白是一种黑色素细胞特异性糖蛋白^[4-5],该位点编码 PMEL17 蛋白,与小鼠的 *silver* 基因同源,编码的蛋白分子质量为 70 ku,由 668 个氨基酸构成;此外,该基因还编码 gp100 蛋白,这 2 种蛋白质通过基因交替剪接成 2 个竞争性 3'接受位点而产生,由于催化活性的不同,这 2 个位点可产生 2 种蛋白质^[6]。在色素细胞中,黑素小体是一种专门用于合成、储存、转运黑素的细胞器,作为哺乳动物和其他脊椎动物的色素沉着来源^[7]。PMEL 作为黑素小体的一种结构蛋白,参与黑素小体由 I 期到 II 期的发育,这对后期黑素小体的形成及黑色素的沉积尤为重要。动物的被毛颜色是重要的质量性状和经济性状,*PMEL* 基因是最优秀的稀释基因家族成员,该基因在鼠^[8]、马^[9]、牛^[10]、猫^[11]、鸡^[12]等家养动物中的突变已被报道,这些突变均能使动物形成银色表型,并逐渐损伤动物的听觉和视觉功能。随着人们对自然风潮的崇尚,天然毛色羊绒和羊毛的选育、生产已有广泛的市场基础,但鲜有关于山羊 *PMEL* 基因结构和功

能的相关研究。本研究在已有山羊 *PMEL* 基因编码蛋白序列的基础上,利用生物信息学方法对其理化性质、疏水性、信号肽、二级结构、亚细胞定位、功能位点及其功能进行预测,基于邻接法对不同物种 *PMEL* 基因进行系统发育分析,为后期研究山羊 *PMEL* 基因结构和功能对其突变的影响提供依据。

1 材料与方法

1.1 材料

通过 NCBI (<http://www.ncbi.nlm.nih.gov/>) 的 Blast 检索人类、小家鼠、家牛等 16 个物种的 *PMEL* 基因蛋白质序列 (表 1)。

表 1 不同物种 *PMEL* 基因编码氨基酸序列

| 物种 | 登录号 | 氨基酸长度(个) |
|--------|----------------|----------|
| 人类 | NP_008859.1 | 661 |
| 小家鼠 | NP_068682.2 | 626 |
| 家牛 | NP_001073684.2 | 649 |
| 原鸡 | NP_990443.2 | 763 |
| 绵羊 | XP_012016119.1 | 647 |
| 家猫 | XP_003988914.3 | 663 |
| 大熊猫 | XP_011219001.1 | 619 |
| 单峰驼 | XP_010975198.1 | 636 |
| 野牛 | XP_010837498.1 | 649 |
| 黑猩猩 | XP_009423449.1 | 589 |
| 苏门答腊猩猩 | XP_009246155.1 | 575 |
| 东非狒狒 | XP_009179187.1 | 661 |
| 褐家鼠 | XP_006240841.1 | 680 |
| 羊驼 | XP_006218617.1 | 644 |
| 山羊 | XP_005680442.1 | 685 |
| 裸鼯鼠 | XP_004861412.1 | 663 |

1.2 方法

1.2.1 山羊 *PMEL* 基因开放阅读框的确定 通过 NCBI 中

收稿日期:2015-06-25

基金项目:河北省高校创新团队领军人才培养计划(编号:LJRC004);河北省应用基础研究计划(编号:15962901D);河北省科技计划(编号:15226302D);河北省现代农业产业技术体系蛋鸡产业创新团队项目。

作者简介:李丽莎(1989—),女,硕士研究生,主要从事动物遗传育种研究。E-mail:15131276262@163.com。

通信作者:李祥龙,博士,教授,博士生导师,主要从事动物遗传育种研究。E-mail:lixianlongen@yahoo.com。

的 ORF finder 确定山羊 *PMEL* 基因开放阅读框。

1.2.2 山羊 *PMEL* 基因编码蛋白的生物信息学分析 通过 ExPASy 数据库 (<http://www.expasy.org/>) 提供的 ProtParam 在线程序预测该基因编码蛋白的氨基酸组成、等电点、正/负电荷残基数、原子总数、摩尔消光系数、半衰期、不稳定指数、脂肪族氨基酸指数、总平均亲水性等;通过 ProtScale 软件预测疏水性。通过 CBS Prediction Servers 数据库 (<http://www.cbs.dtu.dk/services/>) 提供的 SignalP 4.1 Server 在线程序判断该基因编码蛋白是否存在信号肽;通过 TMHMM Server v. 2.0 软件预测跨膜结构区域;通过 Protfun 2.2 Server 软件预测该基因编码蛋白的功能分类。通过 Predict Protein 软件 (<http://www.predictprotein.org/>) 预测蛋白质的二级结构。通过 PSORT II Prediction 软件 (<http://psort.hgc.jp/form2.html>) 预测该蛋白在细胞内的具体存在位置。通过 Motif Scan 软件 (http://myhits.isb-sib.ch/cgi-bin/motif_scan) 预测该基因的功能位点。

1.2.3 不同物种 *PMEL* 基因的系统发育 本研究通过 Clustal W 程序对 16 个物种 *PMEL* 基因编码蛋白序列进行多重序列比对,并利用 MrBayes 软件构建系统发育树。其中,设置原鸡为外群,核酸替代模型为 GTR 模型 ($Nst = 6$),每个位点的替换速率为 Gamma (Rates = Gamma),并运行 100 000 代

(Ngen = 100 000),每 100 代保存 1 棵树 (Samplefreq = 100),舍弃老化样本 (sump burnin = 250) 后再将剩余的样本构建一致树。

2 结果与分析

2.1 *PMEL* 基因的开放阅读框

利用 NCBI 中的 ORF finder 对山羊 *PMEL* 基因的开放阅读框进行定位。起始密码子位于第 14 个核苷酸,终止密码子位于第 2 071 个核苷酸,共编码 685 个氨基酸。

2.2 *PMEL* 基因的理化性质

利用 ProtParam 软件分析山羊 *PMEL* 基因编码蛋白的理化性质。由该蛋白的氨基酸组成 (图 1) 可知,山羊 *PMEL* 基因共编码 685 个氨基酸,其中亮氨酸 (Leu) 含量最高,占全部氨基酸的 10.5%。带负电氨基酸 (Asp + Glu)、带正电氨基酸 (Arg + Lys) 残基的数量分别为 53、51 个。分子式为 $C_{3232}H_{5122}N_{890}O_{984}S_{26}$,原子总数为 10 254。所有半胱氨酸形成胱氨酸 (cystines) 时的消光系数为 98 735 L/(mol · cm),对应的吸光度为 1.352;所有半胱氨酸均不能形成胱氨酸时的消光系数为 97 860 L/(mol · cm),对应的吸光度为 1.340。哺乳动物网织红细胞的半衰期为 30 h,不稳定指数为 46.77,脂肪族氨基酸指数为 84.38,亲水性总平均值为 -0.093。

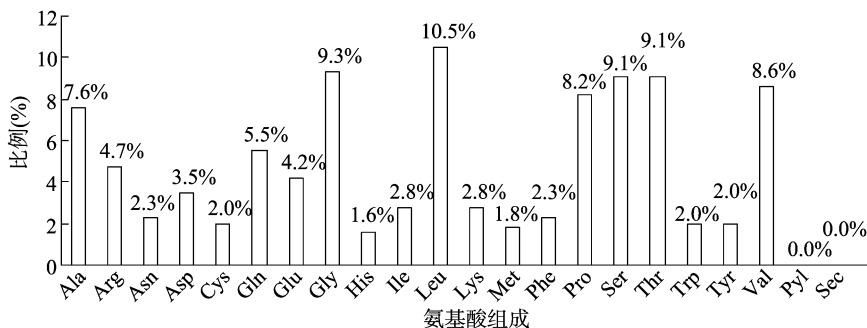


图1 山羊 *PMEL* 基因编码蛋白的氨基酸组成成分

2.3 *PMEL* 基因的疏水性

利用 ProtScale 软件分析山羊 *PMEL* 基因编码蛋白的疏/亲水性 (图 2)。其中,第 624 位异亮氨酸 (I) 得分最高,为 3.389,表示该位点的疏水性最强;第 66 位精氨酸 (R) 得分最低,为 -2.567,表示该位点的亲水性最强。整条多肽链呈现为亲水性,预测山羊 *PMEL* 基因编码蛋白为可溶性蛋白。

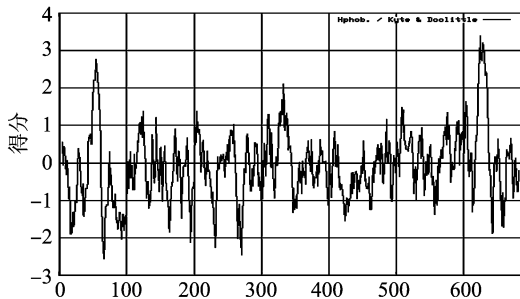


图2 山羊 *PMEL* 基因编码蛋白的疏/亲水性分析结果

2.4 *PMEL* 基因的信号肽

利用 SignalP 4.1 Server 软件分析山羊 *PMEL* 基因编码蛋白的信号肽 (图 3)。结果显示, C 值 (raw cleavage site score,

原始剪切位点得分)、 S 值 (signal peptide score, 信号肽分数)、 Y 值 (combined cleavage site score, 被结合的剪切位点的分数) 不符合 1 个典型信号肽 (C 、 Y 值趋向于 +1; S 值在切割位点之前高,而在切割位点之后降低) 的标准,表明山羊 *PMEL* 基因不存在信号肽。

2.5 *PMEL* 基因的跨膜区

利用 TMHMM Server v. 2.0 软件分析山羊 *PMEL* 基因编码蛋白的跨膜区 (图 4)。结果显示,存在 1 个跨膜区,即 618 ~ 637,表明山羊 *PMEL* 基因编码蛋白为 1 个跨膜蛋白。

2.6 *PMEL* 基因蛋白的功能

通过 Protfun 2.2 Server 软件预测山羊 *PMEL* 基因编码蛋白的功能。结果显示,该蛋白发挥运输和结合、嘌呤和嘧啶、生物合成的辅酶因子、中央中间代谢功能的概率最高,分别为 0.736、0.542、0.135、0.113;在基因本体分类中发挥信号转导、免疫应答作用的概率最高,分别为 0.141、0.111。

2.7 *PMEL* 基因的二级结构

通过 Predict Protein 软件预测山羊 *PMEL* 基因编码蛋白的二级结构。结果显示, H (alpha-helix, α 螺旋): E (beta-sheet, β 折叠): L (random coil, 无规卷曲) = 5.55% : 19.27% : 75.18%,

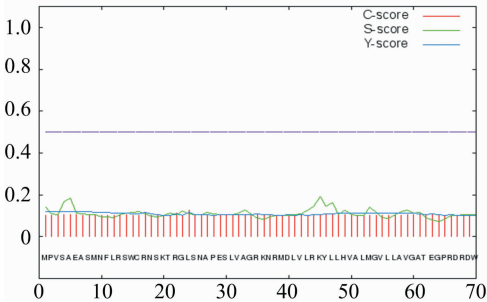


图3 山羊 *PMEL* 基因编码蛋白的信号肽预测结果

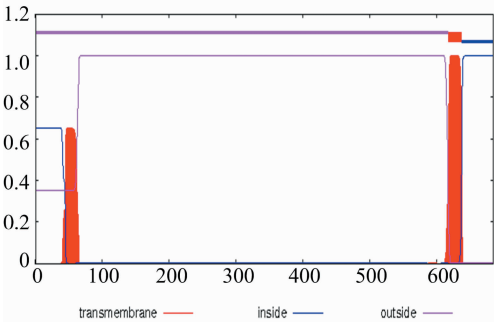


图4 山羊 *PMEL* 基因编码蛋白的跨膜区域预测结果

无规卷曲在整个二级结构中的比例较大,使蛋白质构象呈现出多样性的特性。

2.8 *PMEL* 基因的亚细胞定位

通过 PSORT II Prediction 软件预测山羊 *PMEL* 基因编码蛋白的亚细胞定位。结果显示,该蛋白存在于内质网、线粒体、细胞质膜的概率最高,分别为 34.8%、26.1%、17.4%。

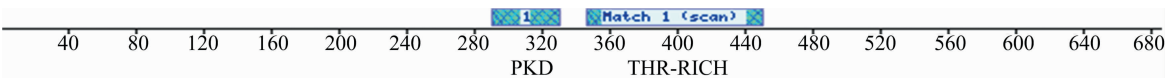


图5 山羊 *PMEL* 基因编码蛋白的功能位点预测结果

2.10 *PMEL* 基因的系统发育分析

通过贝叶斯算法构建的系统发育树(图 6)显示,山羊与绵羊、家牛、野牛的亲缘关系最近,与偶蹄目中单峰驼、羊驼的亲缘关系也较近。食肉目中家猫与大熊猫的亲缘关系最近,灵长目中黑猩猩、东非狒狒、人类、苏门答腊猩猩的亲缘关系

2.9 *PMEL* 基因的功能位点

通过 Motif Scan 软件预测山羊 *PMEL* 基因编码蛋白的功能位点。结果(图 5)显示,多囊肾病区域(polycystic kidney disease domain, PKD)位于第 290 ~ 330 个氨基酸,苏氨酸密集区域(threonine - rich region, THR - RICH)位于第 347 ~ 451 个氨基酸。

最近,啮齿目中小家鼠与褐家鼠的亲缘关系最近。裸鼯鼠虽为啮齿动物,但与小家鼠和褐家鼠的亲缘关系较远,与其他哺乳动物的亲缘关系也较远。哺乳纲与鸟纲(原鸡)之间的亲缘关系最远。上述不同物种间的亲缘关系与生物进化过程中动物学分类相符合。

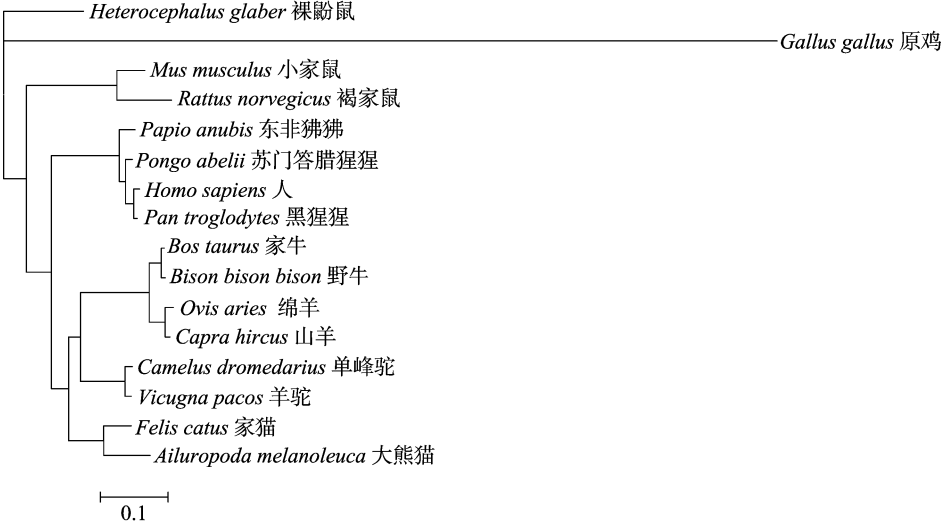


图6 不同物种 *PMEL* 基因的系统发育树

3 讨论

前黑素小体蛋白作为一种黑素细胞特异性糖蛋白,直接参与黑素小体的生物合成,并在黑色素瘤和正常的黑色素细胞中高度表达^[5]。免疫组化(immunocytochemistry)研究结果表明,在不同的毛发生长阶段,黑素细胞表达的蛋白也有差异。当毛发进入生长期,毛球部的黑色细胞具有明显的树突状结构,前黑素小体蛋白、酪氨酸酶、酪氨酸酶相关蛋白 1 均能表达,但在毛囊外根鞘的黑素细胞中只有前黑素小体蛋白

可以表达^[14]。对于前黑素小体蛋白的研究可为黑色素的集聚、黑素细胞的成熟分化提供重要依据。前黑素小体蛋白的突变主要发生在家养动物中,如 Martinez 等发现小鼠 silv 使其体内黑色素细胞丢失,导致毛色变为银灰色^[8]。Brunberg 等发现,该蛋白的多态性与马银色毛色的形成具有极大相关性^[9]。Menotti 等发现,家猫 SILVER 位点的突变对黑色素的产物也有影响^[11]。Schmutz 等发现, *PMEL* 的 1 个缺失和 *PMEL* 与 *MC1R* 之间的互作对能影响高原牛的毛色^[15]。

利用生物信息学的方法对山羊 *PMEL* 基因编码蛋白的氨

基酸组成、疏/亲水性、信号肽、跨膜区等一级结构、二级结构、功能位点、亚细胞定位进行预测。为使预测和分析结果具有可靠性,采用多种不同软件,且预测分析软件所用的原理和算法各不相同,预测和分析结果具有较高的可信度^[16]。推算出山羊 *PMEL* 蛋白由 685 个氨基酸组成,其中亮氨酸含量最高,整条多肽链呈现为亲水性,无信号肽。该基因编码蛋白是一种跨膜蛋白,存在 1 个跨膜区,并包含 1 个 PKD 区和苏氨酸富集区,这与 Adema 等、Maresh 等对前黑素小体蛋白的研究结果^[17-18]一致。已有研究表明,前黑素小体蛋白 17 包含 1 个 RPT 重复区域 (proline/serine/threonine - rich repeat domain),该区域可决定体内黑素小体丝状假足的形成,并影响黑色素的合成功率^[13]。本研究预测出的山羊 *PMEL* 基因编码蛋白的 1 个苏氨酸富集区,可能对山羊绒(毛)的黑色素合成具有重要影响。山羊前黑素小体蛋白主要在内质网中发挥运输、结合、信号转导作用。该蛋白的二级结构组成中,无规卷曲所占比例较大,使蛋白质的构象呈现出多样性的特性。利用贝叶斯算法对 *PMEL* 基因进行系统发育分析,发现山羊 *PMEL* 基因与绵羊、家牛、野牛的亲缘关系最近,哺乳纲与鸟纲之间的亲缘关系最远,这不仅符合动物学分类,并反映了不同物种 *PMEL* 编码蛋白结构上的稳定性对其生物学功能的重要意义。

4 结论

山羊 *PMEL* 蛋白由 685 个氨基酸组成,是一种不稳定的可溶性蛋白,无信号肽,在接近肽链的 C - 端存在 1 个跨膜区。该蛋白主要在内质网中发挥运输、结合、信号转导作用。在二级结构中,无规卷曲所占比例较大,使蛋白质的构象呈现出多样性的特性。系统发育分析表明,*PMEL* 基因在生物进化过程中符合动物学分类,具有较高的保守性。

参考文献:

- [1] Watt B, Raposo G, Marks M S. PMEL17: an amyloid determinant of organelle structure [M] // Rigacci S, Bucciantini M. Functional amyloid aggregation. Trivandrum, Kerala, India: Research Signpost, 2010: 89 - 113.
- [2] Brouwenstijn N, Slager E H, Bakker A B, et al. Transcription of the gene encoding melanoma - associated antigen gp100 in tissues and cell lines other than those of the melanocytic lineage [J]. British Journal of Cancer, 1997, 76(12): 1562 - 1566.
- [3] Leonhardt R M, Vigneron N, Hee J S, et al. Critical residues in the PMEL/Pmel17 N - terminus direct the hierarchical assembly of melanosomal fibrils [J]. Molecular Biology of the Cell, 2013, 24(7): 964 - 981.
- [4] Kwon B S, Halaban R, Kim G S, et al. A melanocyte - specific complementary DNA clone whose expression is inducible by melanotropin and isobutylmethyl xanthine [J]. Molecular Biology & Medicine, 1987, 4(6): 339 - 355.
- [5] Berson J F, Harper D C, Tenza D, et al. Pmel17 initiates premelanosome morphogenesis within multivesicular bodies [J]. Molecular Biology of the Cell, 2001, 12(11): 3451 - 3464.
- [6] Solano F, Martínez E M, Jiménez C C, et al. New insights on the structure of the mouse silver locus and on the function of the silver protein [J]. Pigment Cell Research, 2000, 13(Suppl 8): 118 - 124.
- [7] 呼 蕾, 孙丽君, 李 元, 等. 白癜风患者血清前黑素小体蛋白 PMEL17 自身抗体的检测 [J]. 细胞与分子免疫学杂志, 2013, 29(4): 411 - 413.
- [8] Martinez E M, Jimenez C C, Solano F, et al. The murine silver locus: coding and expression of a single transcript truncated by the silver mutation [J]. Mamm Genome, 1999, 10: 1168 - 1171.
- [9] Brunberg E, Andersson L, Cothran G, et al. A missense mutation in PMEL17 is associated with the silver coat color in the horse [J]. BMC Genetics, 2006, 7: 46.
- [10] Mészáros G, Petautschnig E, Schwarzenbacher H, et al. Genomic regions influencing coat color saturation and facial markings in Fleckvieh cattle [J]. Animal Genetics, 2015, 46(1): 65 - 68.
- [11] Menotti R M, David V A, Eizirik E, et al. Mapping of the domestic cat "Silver" coat color locus identifies a unique genomic location for Silver in mammals [J]. The Journal of Heredity, 2009, 100(Suppl 1): 8 - 13.
- [12] Kerje S, Sharma P, Gunnarsson U, et al. The dominant white, dun and smoky color variants in chicken are associated with insertion/deletion polymorphisms in the *PMEL17* gene [J]. Genetics, 2004, 168(3): 1507 - 1518.
- [13] McGlinchey R P, Shewmaker F, Mcphie P, et al. The repeat domain of the melanosome fibril protein Pmel17 forms the amyloid core promoting melanin synthesis [J]. Proceedings of the National Academy of Sciences of the United States of America, 2009, 106(33): 13731 - 13736.
- [14] Commo S, Bernard B A. Melanocyte subpopulation turnover during the human hair cycle: an immunohistochemical study [J]. Pigment Cell Research, 2000, 13(4): 253 - 259.
- [15] Schmutz S M, Dreger D L. Interaction of MC1R and PMEL alleles on solid coat colors in Highland cattle [J]. Animal Genetics, 2013, 44(1): 9 - 13.
- [16] 王明亮, 陈 鹏, 郝明超, 等. 黄牛 *FSHR* 基因的生物信息学分析 [J]. 甘肃农业大学学报, 2012, 47(4): 7 - 13.
- [17] Adema G J, de Boer A J, Vogel A M, et al. Molecular characterization of the melanocyte lineage - specific antigen gp100 [J]. The Journal of Biological Chemistry, 1994, 269(31): 20126 - 20133.
- [18] Maresh G A, Wang W C, Beam K S, et al. Differential processing and secretion of the melanoma - associated ME20 antigen [J]. Archives of Biochemistry and Biophysics, 1994, 311(1): 95 - 102.