

# 利用错配碱基产生酶切位点快速检测禽类 Mx 蛋白基因单核苷酸改变

尹春光<sup>①②</sup> 李善刚<sup>①</sup> 路国彬<sup>③</sup> 杜立新<sup>①\*</sup>

(① 中国农业科学院北京畜牧兽医研究所 国家畜禽分子遗传育种中心 北京 100094;

② 山东省济宁学院生物学系 济宁 273155; ③ 内蒙古农业大学动物科学与医学学院 呼和浩特 010018)

**摘要:** 该方法针对无合适酶切位点的检测位点设计引物,应用于禽类 Mx 蛋白基因 631 位氨基酸位点的突变检测,利用错配碱基产生内切酶识别的酶切位点检测 SNP,尝试一种操作简便结果可靠的检测单核苷酸多态性并能快速判定 631 位点氨基酸的方法。实验结果表明,利用这种方法产生的酶切位点能有效地对 631 位氨基酸位点进行分型,可以快速检测 SNP。

**关键词:** 错配碱基; 酶切位点; SNP

中图分类号: Q81 文献标识码: A 文章编号: 0250-3263(2008)05-80-06

## An Effective Method for Detecting SNP Using Mismatched Bases to Create Specific Restriction Enzyme Cutting Sites

YIN Chur Guang<sup>①②</sup> LI Shan Gang<sup>①</sup> LU Guo Bin<sup>③</sup> DU Li Xin<sup>①\*</sup>

(① *National Center for Molecular Genetics and Breeding of Livestock and Poultry, Beijing Institute of Animal Science and Veterinary Medicine, Chinese Academy of Agricultural Sciences, Beijing 100094;* ② *Biology Department, Jining University, Jining 273155;* ③ *College of Animal Science and Medicine, Inner Mongolia Agricultural University, Huhhot 010018, China*)

**Abstract:** In order to find a simple and effective way to detect SNP, a method for creation of restriction enzyme cutting sites was developed by using mismatched primers. This method can be used to detect SNP effectively and rapidly.

**Key words:** Mismatched base; Endonuclease sites; SNP

单核苷酸多态性 (single nucleotide polymorphism, SNP) 是指在基因组水平上由单个核苷酸变异所引起的 DNA 序列多态性。SNP 在人类基因组中广泛存在,已经发现了 370 万个 SNP,平均每 500~1000 个碱基对中就有一个。近年来对鸡 (*Gallus gallus*) 基因组 DNA 和表达序列标签 (expressed sequenced tag, EST) 的研究表明,单核苷酸多态性在其中广泛分布。在 8 个欧洲国家实验室合作研究的 AVIANDIV (avian diversity) 项目中发现每 50 bp 就有一个 SNP,表明包括野生红色原鸡在内的被测群体有较宽的遗传图谱<sup>[1,2]</sup>。SNP 是基因控制重要性状的确定、估计以及分析生物多样性和系统

发生关系的有力工具,可用 SNP 等位基因与表型的关系及 SNP 等位基因的变异来提高鸡的育种水平,检测包括重要性状遗传变异在内的基因,阐明其作用和功能,还能将 SNP 用于研究野生及家养群体的遗传结构。通过 PCR 扩增片段的多态性分析有助于检测各种突变,PCR-RFLP 分析、PCR 结合 ASO 探针杂交分析、

基金项目 北京市自然科学基金项目 (No. 5051004);

\* 通讯作者, E-mail: lxdu@263.net;

第一作者介绍 尹春光,女,副教授,博士研究生;研究方向:动物遗传育种与繁殖;E-mail: jnycg@126.com.

收稿日期: 2008-03-28, 修回日期: 2008-07-02

PCR-SSCP 分析、扩增片段长度多态性(AFLP)、PCR 直接测序等技术和荧光 PCR 检测点突变、反向限制性位点突变分析、基因芯片、等位基因特异性扩增技术等新技术的应用使点突变检测技术得到迅速的发展。PCR-RFLP 是检测 SNP 基因型的经典方法, 利用 DNA 序列单核苷酸改变使某种限制性内切酶识别位点出现或消失, 酶切后产生含有相同或者不同长度的片段, 从而判断 DNA 分子单核苷酸的变化。虽然现有的限制性内切酶种类不少, 但还是有很多的 SNP 位点找不到合适的限制性核酸内切酶来识别, 在引物的 3' 末端引入错配碱基产生合适的酶切位点进行 SNP 研究的方法应运而生。在人类 (*Homo sapiens*)<sup>[3]</sup>、羊 (*Ovis aries*)<sup>[4]</sup>、马 (*Equus caballus*)<sup>[5]</sup>、鸡 (*Gallus gallus*)<sup>[6]</sup> 上先后出现了利用这种技术研究单核苷酸改变的方法, 有的称为 forced PCR, 有的称为 mismatched PCR。本文中用“利用错配碱基产生酶切位点”来代表这种方法。

Mx 蛋白是在 IFN 诱导下产生的蛋白质, 是继 2', 5' 寡腺苷酸合成酶和 P86 蛋白激酶后发现的又一种抗病毒蛋白。1962 年, Lindenmann 首次发现小鼠 A2G16 号染色体上有一个显性基因能抵抗正黏液病毒科的病毒(流感病毒), 并将该基因命名为 myxovirusresistant, 即 Mx 基因, 该基因表达使得 A2G 对流感病毒 NWS 株不敏感, 能抵抗致死剂量的流感病毒<sup>[7]</sup>。Ko 等人证实了鸡 Mx 基因具有抗 VSV 病毒、流感病毒的活性, 而且这种活性与第 631 位氨基酸 (Asn/Ser) 有密切关系<sup>[8,9]</sup>, 因此利用鸡的 Mx 基因进行抗病研究具有重大的现实意义。

由于本课题在对禽类 631 位氨基酸对应的核苷酸位点进行研究时, 遇到无合适的内切酶进行 RFLP 的问题。因此引入了错配碱基创造酶切位点检测 SNP 的方法。错配碱基创造酶切位点检测 SNP 的方法是人为在设计的引物上引入错配碱基, 使得引物的 3' 末端与及邻近的几个碱基(包括引物和模板)共同组成一个酶切位点, PCR 产物用相应的限制性内切酶切割, 根据片段长度的差异进行单核苷酸改变的识

别。这种方法使无合适的限制性内切酶进行酶切的 SNP 位点能通过 PCR-RFLP 法进行检测。

## 1 材料与方 法

### 1.1 基因组 DNA 的制备

白莱航和北京油鸡作为实验样本, 每只鸡翅下静脉采血 30  $\mu$ l, 基因组 DNA 的提取参照 Jon 等的方法<sup>[10]</sup>。①收集全血 25  $\mu$ l, 加入 475  $\mu$ l STE buffer (0.15 mol/L NaCl, 0.05 mol/L Tris, 1 mmol/L EDTA pH 8.0) 于 -80  $^{\circ}$ C 保存; ②加入 15  $\mu$ l 蛋白酶 K (10 mg/ml) 和 7.5  $\mu$ l 25% (v/v) SDS 55  $^{\circ}$ C 消化过夜; ③酚, 酚: 氯仿: 异戊醇 (25: 24: 1), 氯仿: 异戊醇 (24: 1) 各一次; ④加入 0.1 倍体积 3 mol/L NaAc, 2 倍体积冰乙醇, 轻轻颠倒混匀, 11 600 g 离心 5 min, 收集絮状 DNA 沉淀; ⑤5% 乙醇洗 2~3 次, 干燥, 加入 150  $\mu$ l TE (10 mmol/L Tris, 1 mmol/L EDTA) 溶解 DNA。

### 1.2 主要试剂和器材

琼脂糖(西班牙)、限制性核酸内切酶(NEB)、蛋白酶 K(Amresco)、DNA 分子量标准(天根)。电泳仪(BIO-RAD)、电泳槽(BIO-RAD)、凝胶成像系统(BIO-RAD)、水浴恒温震荡器(江苏太仓)、PCR 仪(BIO-RAD)。

### 1.3 引物设计

对 GED 区 631 位的碱基序列进行分析(图 1), 其中碱基 A 和 G 的变化在基因组 DNA 的 2 216 位碱基处, 该位点在外显子 14 的起始部分, 因此利用内含子 13 的 3' 末端序列和外显子 14 的 5' 端序列扩增这段区域(图 2), 设计与序列严格互补的上游引物和下游引物(pCF: cctt cagcctgtttttctcttita, pCR: TAAACGCC-TGAGCAATCAGATTCCTC) 验证引物的特异性, 命名为引物 pC。

由于 631 位对应的碱基找不到合适的酶切位点, 此时引入错配碱基, 设计两对下游错配引物 pA 和 pB。其中, pAR: 5'-GAGGAATCTGATT-GCTCAGGCGTgTA-3', 小写的 g 为引入的错配碱基, GTAC 序列是限制性内切酶 *Cvi*Q I 的识别序列, 扩增产物紧靠下游引物 3' 末端的第一个碱基 G 即是要判断的 SNP, 该碱基的改变会导致酶切位点的消失; pBR: 5'-GAGGAATCTG-ATTGCTCAGGCGTgTA-3', 小写的 a 为引入的错

配碱基, AATT 是限制性内切酶 *Tp*509 I 的识别序列, 扩增产物紧靠上游引物末端的第一个碱基 A 即是要判断的 SNP, 该碱基的改变同样

会导致酶切位点的消失(图 3)。利用引物 (pCF, pAR) 及 (pCF, pBR) 进行扩增, 两对引物分别命名为 pA、pB。

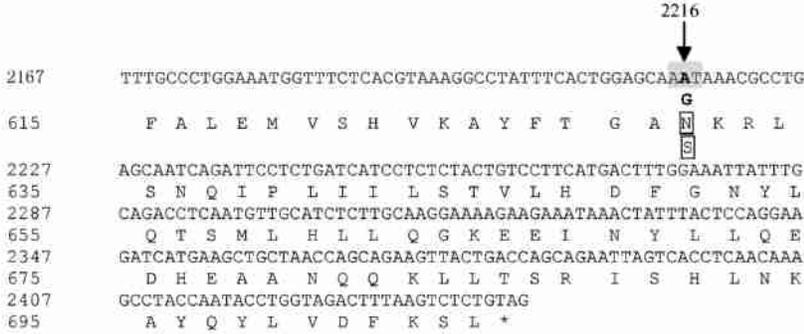


图 1 鸡 Mx 蛋白基因 GED 区序列

Fig. 1 GED sequence of chicken Mx

阴影部分为 Asn 序列编码 AAT, 箭头指示 2 216 位对应的碱基 A 和 G 突变, 对应的 631 位氨基酸 Asn (N) 或 Ser (S) 用方框标明。

Asn codon AAT was shaded, The arrowhead indicates the mutation of G and A at position 2 216, the corresponding amino acid Asn (N) or Ser (S) was boxed.

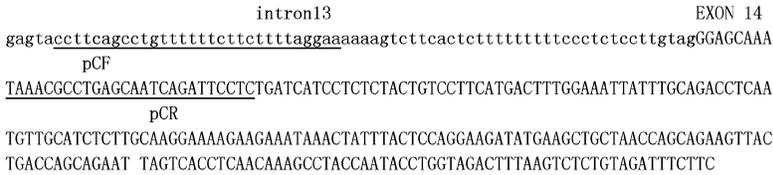


图 2 进行碱基错配设计的区域

Fig. 2 The region used to design mismatched primer

内含子序列用小写字母表示 (intron 13), 外显子序列用大写字母表示 (EXON 14), 引物 pCF、pCR 用下划线标明。

Capital letters indicate nucleotides in exon 14, lowercase letters show nucleotides in intron 13. Primers of pCF, pCR were underlined.

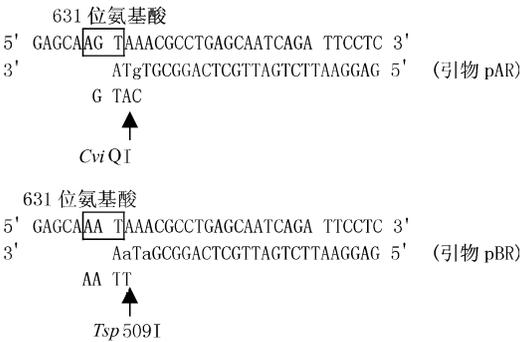


图 3 利用错配引物引入酶 *Cvi*Q I , *Tsp*509 I 的酶切位点

Fig. 3 Creation restriction enzyme cutting site by mismatched primers

1.4 PCR 扩增 PCR 扩增体系为: 基因组 DNA 0.05~ 0.5 μg, dNTPs 2.5 mmol/L, 2.5 μl 10 × PCR 缓冲液, MgCl<sub>2</sub> 1.5 mmol/L, 引物各 20 pmol 和 *Taq* DNA 聚合酶 1 U, 总体积 25 μl。扩增条件: 94 °C 预变性 3 min, 30 个循环 (94 °C 30 s, 60 °C 30 s, 72 °C 60 s), 72 °C 延伸 5 min, 4 °C 保温 5 min。

1.5 酶切 取 6 μl PCR 产物, pA、pB 引物扩增的片段分别用 3 U *Cvi*Q I 及 *Tp*509 I 限制性内切酶酶切, 1.2 μl 10 × 酶切 buffer, 反应总体积 12 μl, 37 °C 酶切过夜。用 3% 琼脂糖凝胶进行电泳, 溴乙锭染色, 100 V/40 min, 凝胶成像系统下观察、记录、照相。

## 2 结果

基因组 DNA 检测结果见图 4, 可见所提取的 DNA 整齐、集中, 无脱尾现象。三对引物 pA、pB、pC, 扩增片段为 101 bp。引物 pA 的产物, 当 631 位 AA 对应的碱基是 G 时, 用 *Cvi*Q I 酶切会出现 75+ 26 bp 的带, 26 bp 的带在 3% 的琼脂糖中看不清, 因此对于纯合的个体出现 101 bp 的带为 AA 型, 出现 75 bp 的带为 BB 型,

杂合的出现 75 bp 和 101 bp 的带为 AB 型。引物 pA 的产物, 当 631 位 AA 对应的碱基是 A 时, 用 *Tsp* 509 I 酶切同样会出现 75+ 26 bp 的带, 27 bp 的带在 3% 的琼脂糖中看不清, 因此对于纯合的个体会出现 75 bp 的带型, 为 AA 型, 杂合的出现 75 bp 和 101 bp 的带, 为 AB 型, 而碱基不是 A 的出现 101 bp 的带, 为 BB 型。因此, 根据使用的引物及相应的酶切就可以辨别出该位点是否有要寻找的抗性位点 G(图 5)。



图 4 鸡基因组 DNA 检测

Fig. 4 Detection of chicken genome DNA

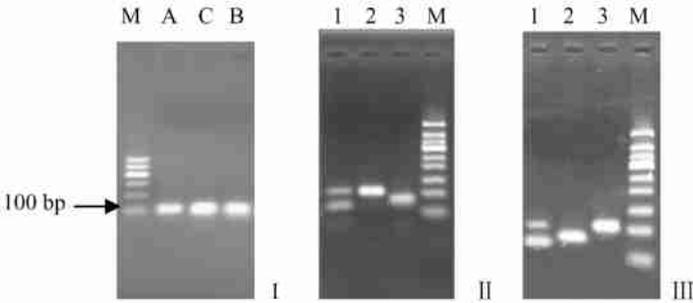


图 5 引物 pA、pB、pC 扩增产物及酶切产物

Fig. 5 Primers A, B and C PCR products and restriction enzyme cutting products

I 为引物 pA、pB、pC PCR 产物: A. pA 扩增 101 bp 带, B. pB 扩增 101 bp 带, C. pC 扩增 101 bp 带, M. DNA 分子量标准; II 为 *Cvi*Q I 酶切引物 pA 扩增片段: 1. 101 bp+ 75 bp (AB), 2. 101 bp (AA), 3. 75 bp (BB), M. 50 bp DNA 分子量标准; III 为 *Tsp* 509 I 酶切引物 pB 扩增片段: 1. 101 bp+ 75 bp (AB), 2. 101 bp (BB), 3. 75 bp (AA), M. 50 bp DNA 分子量标准。  
I : A, B, C are products (101 bp) of primers pA, pB and pC respectively, M. DNA marker I; II : *Cvi*Q I restriction enzyme cutting products (101 bp) of pA, 1. 101 bp+ 75 bp (AB), 2. 101 bp (BB), 3. 75 bp (AA), M. 50 bp DNA ladder; III *Tsp* 509 I restriction enzyme cutting products (101 bp) of pB, 1. 101 bp+ 75 bp (AB), 2. 101 bp (BB), 3. 75 bp (AA), M. 50 bp DNA ladder.

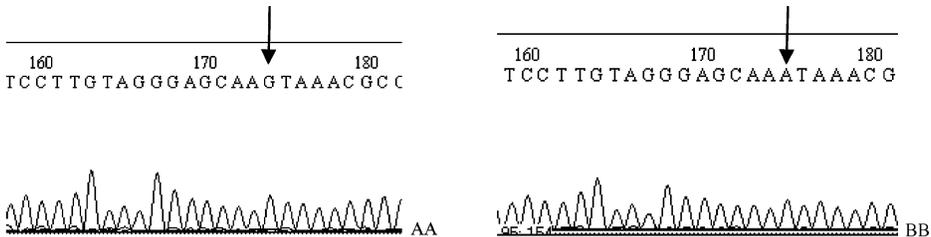


图 6 纯合型 AA 和 BB

Fig. 6 Homozygous of AA and BB

测序验证结果, 送酶切结果为纯合的 DNA 标本到诺赛公司测序, 测序结果与预期一致(图 6)。

用这种方法进行的禽类 *Mx* 基因在 631 位

氨基酸检测结果如图 7, 结果表明用这种方法能够方便快捷地分辨出 631 位位点的多态性, 大大加快了实验过程中对携带抗性位点品种的选种进程。

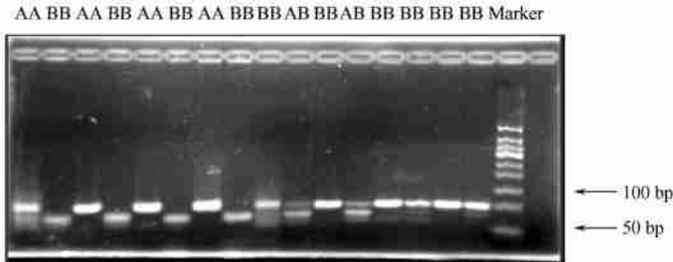


图 7 用错配引物 pA、pB 扩增后获得的 631 位点酶切产物结果  
Fig. 7 Detection of position 631 with mismatched primers (pA, pB)

### 3 讨论

*Mx* 蛋白有独特的结构, 蛋白分子能够自组装成高级的环状和螺旋环状的功能区, 这种结构一般具有以下结构域: 300 个高度保守 N 端 GTPase 区, 含 150 个氨基酸的中部结构域, 大约 100 个氨基酸的 C 端结构域, 该区两个亮氨酸拉链结构域可形成双链  $\alpha$  螺旋, 是 GTPase 反应区(GTPase effect domain, GED)。研究结果显示人类 *MxA* 蛋白 GED 能够特异结合中部区域, 使 *Mx* 蛋白形成二聚体或者三聚体结构, 这种结合对于形成有功能的 GTPase 非常关键<sup>[11, 12]</sup>。尽管这种功能和结构可能会限制不同 *Mx* 蛋白序列变化, 但是 C 端结构域的序列同源性只有 22%, 而其余序列的同源性可达到 38%, C 端结构域这种变化其实与动力蛋白家族中各成员在细胞内的定位及功能有关<sup>[13]</sup>, 因为肽链 C 端有决定 *Mx* 蛋白定位于细胞核还是细胞质的定位信号, 此外 C 末端氨基酸的变化决定了其抗病毒的特异性, 因此可以设想 *Mx* 蛋白在抵抗病毒过程中的功能是 C 末端与病原体直接特异性作用的结果<sup>[14]</sup>。禽类 *Mx* 蛋白抗病毒活性与 GED 区单核苷酸多态性有关, 其中 631 位氨基酸(对应于 cDNA 序列是 2 216 位碱基)位点 N (AAC) 与 S (AGC) 的变化, 即碱基 A 与 G 的变化可以改变其抗病毒的作用<sup>[8, 9]</sup>, 但是在该位点

找不到合适的酶, 引入错配碱基产生酶切位点可以经济、快捷地检测单核苷酸多态性<sup>[15]</sup>。由于在育种过程中首先选择具有抗性位点的碱基 A, 因此在大规模的筛选育种过程中利用该方法方便、经济、快捷, 有利于育种过程中的初步筛选。

传统的 PCR-RFLP 方法在酶的选择上只能严格寻找在序列上完全匹配的有识别序列的限制性核酸内切酶, 然而并非所有的序列都有合适的酶的识别位点, 因此该方法在使用时受到酶的限制。在对 631 位氨基酸进行 SNP 研究时就找不到合适的酶。利用错配碱基在引物设计时在一条引物中人为引入错配碱基产生适合酶切的合适位点, 就能比较方便地进行 SNP 研究。应用错配碱基引入酶切位点的方法进行 SNP 研究, 对 SNP 位点进行分型, 结果可靠、操作简便、经济快捷, 普通分子实验室都可以进行。实验中 631 位点处本身没有酶切位点, 设计错配引物是为了创造一个酶切位点用 PCR-RFLP 对其进行分型。在设计错配引物时, 错配的碱基最好在倒数第二个或稍前处, 错配碱基也不宜过多, 否则影响扩增效率, 同时应该设计一对与原序列互补的引物验证在非错配情况下该引物的特异性, 在保证特异性的基础上才能进行错配引物的 PCR。

根据本实验, 在 PCR 的程序设定上, 由于

加入了错配引物, 反应的特异性相对降低, 在用错配引物而使用原引物扩增程序扩增效率不高时, 可以适当降低退火温度, 或者适当延长退火时间。通过巢式 PCR 二次扩增出该序列, 也会提高扩增效率, 但是必须对一次 PCR 的产物进行浓度梯度稀释, 寻找合适的模板浓度进行第二次 PCR。进行检测时可使用琼脂糖凝胶电泳, 无须进行 PAGE 胶检测。由于酶切后不同长度片段的差异基本上是引物的长度, 为了适于普通的琼脂糖电泳分辨, 错配引物的扩增产物最好在 101 bp 左右, 并使用较高浓度的琼脂糖凝胶 (3%~4%), 可以得到较满意的结果。同时, 为了更好地观察酶切产物的差异, 电泳时间要较长。

## 参 考 文 献

- [ 1 ] Gibbs R A, Belmont J W, Hardenbol P, *et al.* The International HapMap Project. *Nature*, 2003, **426**(6 968): 789~ 796.
- [ 2 ] Altshuler D, Brooks L D, Chakravarti A, *et al.* A haplotype map of the human genome. *Nature*, 2005, **437**(7 063): 1 299~ 13 20.
- [ 3 ] Murray P. Minimal hierarchical analysis of global human Y-chromosome SNP diversity by PCR-RFLP. *Anthropological Science*, 2006, **114**: 69~ 74.
- [ 4 ] Guan F, Liu S R, Shi G Q, *et al.* Polymorphism of *FecB* gene in nine sheep breeds or strains and its effects on litter size, lamb growth and development. *Acta Genetica Sinica*, 2006, **33**(2): 117~ 124.
- [ 5 ] Ju L H, Onogi A, Ueda J, *et al.* Polymorphic study of equine antiviral *MxA* gene. *Biochem Genet*, 2005, **43**: 299~ 305.
- [ 6 ] Seyama T, Ko J H, Ohe M, *et al.* Population research of genetic polymorphism at amino acid position 631 in chicken *Mx* protein with differential antiviral activity. *Biochem Genet*, 2006, **44**: 437~ 448.
- [ 7 ] Lindenmann J. Resistance of mice to mouse adapted influenza A virus. *Virology*, 1962, **16**: 203~ 204.
- [ 8 ] Ko J H, Takada A, Mitsuhashi T, *et al.* Native antiviral specificity of chicken *Mx* protein depends on amino acid variation at position 631. *Anim Genet*, 2004, **35**(2): 119~ 122.
- [ 9 ] Seyama, Ko J H, Ohe M, *et al.* Population research of genetic polymorphism at amino acid position 631 in chicken *Mx* protein with differential Antiviral Activity. *Biochemical Genetics*, 2006, **44**(9 10): 432~ 443.
- [ 10 ] Jon H, Wetton, Royston E, *et al.* Demographic study of a wild house sparrow population DNA fingerprinting. *Nature*, 1987, **327**(14): 147~ 149.
- [ 11 ] Schumacher B, Staeheli P. Domains mediating intramolecular folding and oligomerization of *MxA* GTPase. *J Biol Chem*, 1998, **273**(43): 28 365~ 28 370.
- [ 12 ] Di Paolo C, Hefti H P, Meli M, *et al.* Intramolecular backfolding of the carboxyl terminal end of *MxA* protein is a prerequisite for its oligomerization. *J Biol Chem*, 1999, **274**(45): 32 071~ 32 078.
- [ 13 ] Floyd S, De Camilli P. Endocytosis proteins and cancer: a potential link? *Trends Cell Biol*, 1998, **8**(8): 299~ 301.
- [ 14 ] 尹春光, 杜立新. 禽类抗病毒蛋白质 Mx 蛋白. 生命的化学, 2007, **27**(6): 508~ 510.
- [ 15 ] Lee C Z, Lee H S, Huang G T, *et al.* Detection of YMDD mutation using mutant specific primers in chronic hepatitis B patients before and after lamivudine treatment. *World J Gastroenterol*, 2006, **12**(33): 5 301~ 5 305.