

GUANGXI  
ZHUANGZU RENQUN  
JINYIN DUOTAI XING CHUTAN

广西

壮族人群

基因

多态性

初探

解继胜 黄海玲 主编

广西科学技术出版社



主 编 解继胜 黄海玲  
副主编 韦叶生 刘运广 黄瑞雅

广西科学技术出版社

图书在版编目 (CIP) 数据

广西壮族人群基因多态性初探/解继胜等主编. —南宁:  
广西科学技术出版社, 2007.5

ISBN 978-7-80666-913-6

I. 广... II. 解... III. 壮族—人类基因—多态现象—研  
究—广西 IV. Q987

中国版本图书馆CIP数据核字 (2007) 第 038059 号

广西壮族人群基因多态性初探

解继胜 黄海玲 主编

\*

广西科学技术出版社出版

(南宁市东葛路 66 号 邮政编码 530022)

广西新华书店发行

广西地质印刷厂印刷

(南宁市建政东路 88 号 邮政编码 530023)

\*

开本 890mm×1240mm 1/32 印张 5.25 字数 125 000

2007 年 5 月第 1 版 2007 年 5 月第 1 次印刷

印数: 1—3 000 册

ISBN 978-7-80666-913-6/R·149 定价: 11.00 元

本书如有倒装缺页, 请与承印厂调换

## 序 言

人类是一个具有多态性的群体。不同的群体和个体在生物性状以及在对疾病的易感性/抗性上的差别,反映了进化过程中基因组与内、外环境相互作用的结果。已知人类基因组 DNA 序列中最常见的变异形式是单个碱基多态性 (Single nucleotide polymorphism, SNP), 这种多态性包括单个碱基的缺失和插入以及单个碱基的置换。SNP 作为一种碱基的替换, 大多数为转换, 即一种嘧啶碱基换为另一种嘧啶碱基或一种嘌呤碱基换为另一种嘌呤碱基; 少数为颠换, 即嘧啶碱基与嘌呤碱基之间的替换。转换与颠换之比为 2:1。SNP 在 CG 序列上出现最为频繁, 而且多是 C→T, 原因是 CG 中胞嘧啶常为甲基化的、自发地脱氨后即成为胸腺嘧啶。SNP 为数众多, 分布广泛, 在全基因组中估计有  $3 \times 10^6 \sim 6 \times 10^6$  个。位于基因的编码区 SNP 称为 cSNP。若 cSNP 引起蛋白质重要部位氨基酸的变异, 可导致其功能改变; 位于基因调控序列中的 SNP 则可能影响基因表达的剂量。故这两种 SNP 的生物学意义最为重要, 是基因组中决定人类表型多样性的核心信息。

开展人类基因组多态性的系统研究, 无论对于生物医学, 还是对于了解人类的起源、进化和迁徙均会产生重大的影响, 主要体现在以下几方面。

一、在疾病基因组学, 已知多基因疾病是由多个微效基因的累加作用和某些环境因子作用所致, 估计参与多基因疾病发病原理的基因可以达 3~20 个, 这些基因的 SNP 及其特定组合可能是

造成疾病易感性最重要的原因。

二、在药物基因组学，同一种药物在不同个体产生的效果不是完全相同，这种不同是由于药物本身在不同个体体内活化、代谢、清除方面的差异所决定的，而这种差异首先是遗传差异。基因组的多态性，尤其是 SNP 多态性能充分地反映个体间的遗传差异。通过研究遗传多态性与个体对药物敏感性或耐受性的相关性，可以阐明遗传因素对药物效用的影响，从而对针对性的用药和药物开发提供指导和依据。

三、在人类学，DNA 是生物进化的重要记录者，比较研究相应 DNA 区段的变异速度和序列间差异，人们可以获取丰富的进化信息，为人类的起源、进化和人群的迁移提供可靠的遗传学证据。构建一张遗传标记图谱对解决人类进化中的问题尤为重要，而 SNP 已成为最为普遍的标记。

我国是一个人口大国，而且是一个多民族的群体。我国丰富的人群遗传资源是研究人类基因组多样性、人类疾病相关基因以及人类进化的宝贵材料。广西是我国壮族聚居最集中的地区，我们课题组应用聚合酶链反应 (PCR) 和聚合酶链反应—限制性片段长度多态性 (PCR-RFLP) 的方法获得了广西壮族人群细胞因子 (白细胞介素-1、白细胞介素-6、转化生长因子- $\beta_1$  和肿瘤坏死因子- $\alpha$ )、E-选择素、雌激素受体的基因多态性分布情况，并与国外其他种族人群的研究结果相比较，探讨不同种族和地域间这些基因多态性分布及等位基因频率的异同，同时，初步探讨这些基因多态性与疾病的关联。我们的研究成果丰富了分子人类学的知识体系，也将有助于进一步的群体遗传学研究和基因多态性普查，可为今后在我国壮族人群中研究这些基因多态性与疾病间的相互关系提供实验依据。

本研究获广西自然科学基金 (桂科自 0499004)、广西青年科学基金 (桂科青 0447060)、广西壮族自治区教育厅课题 (桂科教

200420)、右江民族医学院科研基金(右医院字 200486)资助。

由于时间仓促,加之我们水平有限,经验不足,书中难免存在错漏或不足之处,恳请各界同仁和读者批评指正。本书在编印过程中得到了右江民族医学院科研处和组织胚胎学教研室全体教师的大力支持,在此一并致谢。

编著者

2006 年 4 月 18 日

# 目 录

## 第一部分 白细胞介素-1 基因多态性 ..... (1)

一、广西壮族人群 IL-1 $\alpha$  (-889) 与  $\beta$  (-551) 基因多态性分布 ..... (2)

二、广西壮族与汉族人群 IL-1 $\beta$  (+3953) 基因多态性分布 ..... (7)

三、广西壮族人群 IL-1Ra (VNTR) 基因多态性分布 ..... (10)

四、广西壮族人群 IL-1Ra (+8006) 基因多态性分布 ..... (14)

## 第二部分 白细胞介素-6 基因多态性 ..... (20)

一、广西壮族与汉族人群 IL-6 (-634) 基因多态性分布 ..... (20)

二、广西壮族与汉族人群 IL-6 (-572) 基因多态性分布 ..... (23)

## 第三部分 雌激素受体基因多态性 ..... (28)

一、广西壮族与汉族人群 ER- $\alpha$  基因多态性分布 ..... (28)

二、广西壮族与汉族人群 ER- $\beta$  基因多态性分布 ..... (32)

第四部分 E-选择素基因多态性 ..... (38)

- 一、广西壮族与汉族人群 E-选择素 (+98) 基因多态性分布 ..... (38)
- 二、广西壮族与汉族人群 E-选择素 (+128) 基因多态性分布 ..... (42)

第五部分 基因多态性与疾病关联 ..... (48)

(1) 一、IL-1 基因多态性与疾病关联 ..... (48)

- (S) (一) IL-1 $\alpha$  和 IL-1Ra 基因多态性与 2 型糖尿病的关联 ..... (48)

- (二) IL-1 $\beta$  基因多态性与 2 型糖尿病的关联 ..... (55)

- (7) (三) IL-1 基因多态性与原发性高血压的关联 ..... (60)

二、IL-6 基因多态性与疾病关联 ..... (68)

- (101) (一) IL-6 基因多态性与急性心肌梗死的关联 ..... (68)

- (二) IL-6 基因多态性与冠心病的关联性及其对血脂的影响 ..... (75)

- (89) (三) IL-6 基因多态性与脑梗死的关联 ..... (81)

- (四) IL-6 基因多态性与血脂水平及血小板计数的关联 ..... (88)

(09) 三、雌激素受体基因多态性与疾病关联 ..... (93)

- (83) (一) ER- $\alpha$  基因多态性与子宫内膜异位症的关联 ..... (94)

- (83) (二) ER- $\beta$  基因多态性与子宫内膜异位症的关联 ..... (99)

(83) 四、E-选择素基因多态性与疾病关联 ..... (104)

- (88) (一) E-选择素基因多态性与脑梗死的关联 ..... (104)

- (二) E-选择素基因多态性与系统性红斑狼疮的关



|                                           |       |
|-------------------------------------------|-------|
| 联 .....                                   | (110) |
| (三) E-选择素基因多态性与缺血性脑卒中的关联<br>.....         | (115) |
| 第六部分 其他 .....                             | (123) |
| 一、广西壮族与汉族人群 TGF- $\beta_1$ 基因多态性分布 .....  | (123) |
| 二、TGF- $\beta_1$ 基因多态性与 2 型糖尿病肾病的关联 ..... | (127) |
| 三、广西汉族人群 TNF- $\alpha$ 基因多态性分布 .....      | (138) |
| 四、白细胞介素-6 基因多态性的研究进展 .....                | (142) |

## 第一部分 白细胞介素-1 基因多态性

众所周知,随着人类基因组测序工作的完成,人类基因突变位点的筛选、检测及其功能的研究正成为研究者们广泛关注的焦点。研究表明,存在于基因编码区和基因调节区域的某些基因突变能够导致基因功能改变,对表型产生很大的影响,而且在不同人群中,其基因型分布频率的差异,可以代表某一种族或人群间的遗传差异。

白细胞介素-1 (Interleukin-1, IL-1) 是一种具有广泛生物学活性的多肽物质,为细胞因子家族中的一员,可由多种细胞合成和分泌,不但在机体的免疫应答中起重要的调节作用,而且还广泛作用于机体的其他组织和器官,参与细胞的激活、增殖和分化。IL-1 由三种成分组成: IL-1 $\alpha$ 、IL-1 $\beta$  和 IL-1Ra, 前两者是 IL-1 家族的激动剂,而后者是 IL-1 家族的拮抗剂,三者的编码基因依次为 IL-1A、IL-1B 和 IL-1RN。近年来研究表明,在 IL-1 $\alpha$  基因启动子区-889 位点、IL-1 $\beta$  基因启动子区-551 位点、IL-1 $\beta$  基因第 5 号外显子区+3953 位点均存在一种由 C $\rightarrow$ T 转换所引起的突变,在 IL-1Ra 基因第 2 号外显子区+8006 位点存在一种由 T $\rightarrow$ C 转换所引起的突变,而在 IL-1Ra 基因第 2 号内含子区则存在一种数目可变的串联重复序列数 (Variable number of tandem repeat, VNTR) 所发生的突变,上述突变可能影响 IL-1 各成分的表达水平,不同的 IL-1 基因型有可能决定不同个体间 IL-1 的表达水平。IL-1 存在基因多态性,广西是我国壮族聚居最集中的地区,IL-1

基因型在壮族的分布特点如何,国内外均无相关报道。我们针对这一空白进行研究,研究的主要内容是应用限制性片段长度多态性(restriction fragment length polymorphisms, RFLP)和聚合酶链反应(polymerase chain reaction, PCR)方法检测广西壮族人群 IL-1 家族突变位点基因型,并结合国外文献进行了不同种族间的比较。其目的是为了解 IL-1 家族基因型在中国广西壮族人群中的分布特征,丰富分子人类学知识体系,为进一步的群体遗传学研究和基因多态性普查提供资料,为今后在我国壮族人群中研究 IL-1 基因多态性与疾病间的关联提供正常人群基因背景。

## 一、广西壮族人群 IL-1 $\alpha$ (-889) 与 IL-1 $\beta$ (-551) 基因多态性分布

### 1 对象和方法

#### 1.1 实验对象

壮族人 148 名,男 78 名,女 70 名,平均年龄  $38 \pm 10$  岁,全部研究对象均来自广西地区壮族健康人群,且无血缘关系。

#### 1.2 引物设计与合成

参照文献设计两对引物,由北京赛百盛有限公司合成。分别扩增包含 IL-1 $\alpha$  (-889) 和 IL-1 $\beta$  (-551) 位点的 DNA 片段。引物序列分别为: IL-1 $\alpha$  (-889) P1: 5'-AAGCTTGTCTACCACCTGAAGTAGGC-3', P2: 5'-TTACATATGAGCCTTCCATG-3', PCR 产物片段长度为 104 bp; IL-1 $\beta$  (-551) P1: 5'-TGGCATTGATCTGGTTCATC-3', P2: 5'-GTTTAGGAATCTTCCCACTT-3', PCR 产物片段长度为 304 bp。

### 1.3 人类基因组 DNA 提取

参照已建立的改良碘化钠法。取全血 100  $\mu\text{l}$  加无菌双蒸水 100  $\mu\text{l}$  混匀, 加 6 mol/L NaI 200  $\mu\text{l}$  涡旋振荡 30 s, 加等体积氯仿/异戊醇 (24:1) 涡旋振荡 30 s, 离心后取上清液加 0.6 倍体积异丙醇混匀, 室温 (或 4℃) 下静置 15 min, 1500r/min 离心。沉淀以 37% 异丙醇 (或 70% 乙醇) 洗 1 次, 晾干后溶于 TE 中, 所提 DNA 经紫外分光光度计定量后, 置 -20℃ 下保存备用。

### 1.4 聚合酶链反应基因扩增

PCR 扩增反应体系均为 25  $\mu\text{l}$ , 其中含 10  $\times$  PCR Buffer 2.5  $\mu\text{l}$ , 2.5 mmol/L dNTPs 2.0  $\mu\text{l}$ , 引物 P1、P2 各 20 pmol, 模板 DNA 为 200 ng, *Taq* DNA 聚合酶 1.25 U, 不足体积用灭菌双蒸水补足至 25  $\mu\text{l}$ 。置热循环仪 (TC-48/H) 中 94℃ 预变性 5 min; 再按下列程序循环 35 次, 即 94℃ 变性 40 s, 57℃ 退火 1 min, 72℃ 延伸 1 min; 末次循环后, 72℃ 延伸 5 min。

### 1.5 扩增产物的限制性酶切

酶切体系包括 IL-1 $\alpha$  (-889) 和 IL-1 $\beta$  (-551) 位点扩增产物各 10  $\mu\text{l}$ , 分别用限制性内切酶 *Nco* I 12.0 U 和 *Ava* I 6 U 于 37℃ 水浴酶切 3 h。

### 1.6 基因型鉴定

IL-1 $\alpha$  (-889) 和 IL-1 $\beta$  (-551) 位点酶切产物分别用 2% 琼脂糖凝胶电泳-溴化乙锭染色法和 8% 聚丙烯酰胺凝胶电泳-溴化乙锭染色法检测, DNA 标准物分别为 pGEM3Zf (+) DNA/*Hae* III 和 DL 2000 Marker。电泳、染色后在紫外灯下鉴定结果, 并拍照。

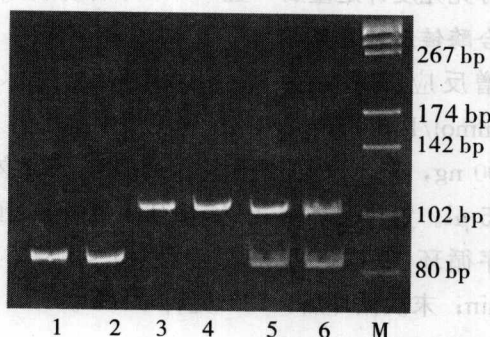
### 1.7 统计学处理

根据正常人群 IL-1 $\alpha$  (-889) 位点和 IL-1 $\beta$  (-551) 位点基因型分布分别计算其等位基因频率, 然后经 Hardy-Weinberg 遗传平衡定律检验。各等位基因及基因型频率用  $\chi^2$  检验。

## 2 结果

### 2.1 IL-1 基因型检测结果

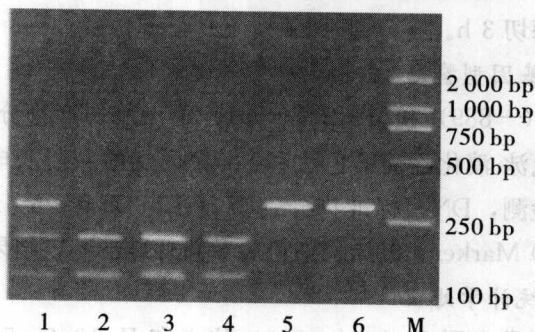
IL-1 $\alpha$  (-889) 多态性分为 CC (88 bp 和 16 bp)、TT (104 bp) 和 CT (88 bp、16 bp 和 104 bp) 三种基因型 (图 1-1); IL-1 $\beta$  (-551) 多态性分为 CC (190 bp 和 114 bp)、CT (190 bp、114 bp 和 304 bp) 和 TT (304 bp) 三种基因型 (图 1-2)。



1, 2: CC 基因型; 3, 4: TT 基因型;

5, 6: CT 基因型; M: pGEM3Zf (+) DNA/HaeIII

图 1-1 IL-1 $\alpha$  基因 -889 位点多态性电泳图



1: CT 基因型; 2, 3, 4: CC 基因型;

5, 6: TT 基因型; M: DL-2000 Marker

图 1-2 IL-1 $\beta$  基因 -551 位点多态性电泳图

2.2 IL-1 $\alpha$  (-889) 和 IL-1 $\beta$  (-551) 位点基因多态性在壮族人群中的分布

广西壮族健康人群中 IL-1 $\alpha$  (-889) 位点基因型以 CC 纯合子型最为多见, CT 杂合子型次之, TT 纯合子型罕见; 其等位基因的频率分布以 C 等位基因最为多见, 其次为 T 等位基因。IL-1 $\beta$  (-551) 位点基因型以 CT 杂合子型最为多见, CC 纯合子型次之, TT 纯合子型最少; 其等位基因的频率分布也是以 C 等位基因最为多见, 其次为 T 等位基因 (见表 1-1)。

2.3 不同种族间 IL-1 $\alpha$  (-889) 和 IL-1 $\beta$  (-551) 位点基因多态性分布的比较

将本实验检测结果分别与奥地利人、非洲黑人、非洲白人及日本人的健康人群比较, 发现在中国广西壮族健康人群中 IL-1 $\alpha$  (-889) 分布频率明显不同于奥地利人、非洲黑人及非洲白人 ( $P < 0.001$ ), 与日本人群接近 ( $P > 0.05$ ); IL-1 $\beta$  (-551) 分布频率明显不同于奥地利人、非洲白人 ( $P < 0.001$ ), 与非洲黑人、日本人群接近 ( $P > 0.05$ )。

表 1-1 壮族与其他种族间基因多态性分布的比较

| 基因            | 种族     | 例数  | 基因型频率(%)  |           |          | 等位基因频率(%) |           |
|---------------|--------|-----|-----------|-----------|----------|-----------|-----------|
|               |        |     | CC        | CT        | TT       | C         | T         |
| IL-1 $\alpha$ | 壮族     | 148 | 133(89.9) | 12(8.1)   | 3(2.0)   | 278(93.9) | 18(6.1)   |
|               | 奥地利人** | 228 | 101(44.3) | 114(50.0) | 13(5.7)  | 316(69.3) | 140(30.7) |
|               | 非洲黑人** | 73  | 18(24.7)  | 43(58.9)  | 12(16.4) | 79(54.1)  | 67(45.9)  |
|               | 非洲白人** | 202 | 68(33.7)  | 109(54.0) | 25(12.3) | 245(60.6) | 159(39.4) |
|               | 日本人    | 112 | 105(93.8) | 7(6.2)    | 0(0.0)   | 217(96.9) | 7(3.1)    |
| IL-1 $\beta$  | 壮族     | 148 | 45(30.4)  | 68(45.9)  | 35(23.7) | 158(53.4) | 138(46.6) |
|               | 奥地利人*  | 227 | 88(38.8)  | 112(49.3) | 27(11.9) | 288(63.4) | 166(36.6) |
|               | 非洲黑人   | 73  | 21(28.8)  | 27(37.0)  | 25(34.2) | 69(47.3)  | 77(52.7)  |
|               | 非洲白人*  | 202 | 89(44.1)  | 87(43.1)  | 26(12.8) | 265(65.6) | 139(34.4) |
|               | 日本人    | 112 | 31(27.7)  | 58(51.8)  | 23(20.5) | 120(53.6) | 104(46.4) |

与壮族人群比较, \*  $P < 0.05$ , \*\*  $P < 0.001$

### 3 讨论

人类 IL-1 基因簇定位于 2 号染色体长臂 1 区 3 带, IL-1 $\alpha$  和 IL-1 $\beta$  基因结构相似, 其基因启动子区域存在一个单核苷酸多态性, 由 C $\rightarrow$ T 转换所引起。近来研究表明, IL-1 $\alpha$  (-889) 和 IL-1 $\beta$  (-551) 位点基因多态性与胃癌、冠心病、II 型糖尿病等多种疾病密切相关, 而且不同种族间基因多态性的分布可能存在着差异。要从基因水平上分析和阐明不同疾病在不同种族、不同民族间患病率的差异与发病机制, 首先要了解正常群体中基因多态性的分布特征及种族差异。广西是我国壮族聚居地区, 我们应用聚合酶链反应限制性片段长度多态性 (PCR-RFLP) 的方法首次获得了广西壮族 IL-1 $\alpha$  (-889) 和 IL-1 $\beta$  (-551) 位点基因多态性的分布情况。同时, 从我们的结果可以看出, IL-1 $\alpha$  (-889) 和 IL-1 $\beta$  (-551) 位点基因多态性分布及等位基因频率在不同种族间存在差异, 造成此结果的可能原因是由于不同民族及种族间基因的遗传背景不同所引起。

总之, 对 IL-1 $\alpha$  和 IL-1 $\beta$  基因多态性的研究将有助于进一步认识与其相关的多种疾病在不同种族间所表现出来的差异。此外, 本实验得到有关 IL-1 基因多态性在我国壮族健康人群中的分布情况, 也将有助于进一步的群体遗传学研究和基因多态性普查, 可为今后在我国壮族人群中研究 IL-1 基因多态性与疾病间的相互关系提供实验依据。

(解继胜 黄海玲)



## 二、广西壮族与汉族人群 IL-1 $\beta$ (+3953) 基因多态性分布

### 1 对象与方法

#### 1.1 研究对象

广西地区壮族人 155 名,男 80 名,女 75 名,平均年龄  $36.9 \pm 9.3$  岁;广西地区汉族人 195 名,男 100 名,女 95 名,平均年龄  $(37.6 \pm 8.7)$  岁,血常规、血脂及其他生化指标均在参考范围内,心电图检查正常,排除肝脏、肾脏、内分泌和心脑血管疾病。全部研究对象均来自健康人群,且无血缘关系。

#### 1.2 研究方法

1.2.1 模板 DNA 的制备 采集体检健康者静脉血 2 ml,用 EDTA-K<sub>2</sub> 抗凝;采用已建立的改良碘化钠法提取白细胞基因组 DNA,置  $-30^{\circ}\text{C}$  下保存备用。

1.2.2 引物合成 参照文献设计合成一对引物(北京赛百盛有限公司合成),扩增 IL-1 $\beta$  基因外显子区的一段 DNA 序列。上游引物:5'-GTTGTCATCAGACTTTGACC-3',下游引物:5'-TTCAGTTCATATGGACCAGA-3'。

1.2.3 PCR 扩增条件 PCR 反应体系包括 25 ng 基因组 DNA,2.5  $\mu\text{l}$  10  $\times$  PCR Buffer,2.0  $\mu\text{l}$  2.5 mmol/L dNTPs,1.5  $\mu\text{l}$  25 mmol/L MgCl<sub>2</sub>,引物各 20 pmol,Taq DNA 聚合酶 1U,所有试剂均为 Promega 产品。PCR 反应循环参数:95 $^{\circ}\text{C}$  预变性反应 4min;95 $^{\circ}\text{C}$  变性 1 min,59 $^{\circ}\text{C}$  退火 1 min,72 $^{\circ}\text{C}$  延伸 2 min,33 个循环;72 $^{\circ}\text{C}$  延伸 4 min,4 $^{\circ}\text{C}$  保存,以上反应均在 GeneAmp PCR System 2700 型(美国 PE 公司)基因扩增仪上完成。



1.2.4 扩增产物的限制性酶切 取 PCR 扩增产物 10  $\mu$ l, 用 15U *Taq* I 于 37 $^{\circ}$ C 下酶切 3 h。反应终止后, 产物经 3% 琼脂糖凝胶电泳, EB 染色, 染色后以 DL2000 DNA 片段长度标准物为参考, 在紫外灯下判断结果, 并拍照。

### 1.3 统计学处理

用基因直接计数法计算健康人群 IL-1 $\beta$  的基因型及等位基因频率, 然后经 Hardy-Weinberg 遗传平衡定律检验。不同人群间基因型及等位基因频率用  $\chi^2$  检验, 均在 SPSS (statistical package for the social science) 11.5 软件包上完成。以  $P < 0.05$  为具有显著性差异。

## 2 结果

### 2.1 IL-1 $\beta$ (+3953) 位点基因型分析

IL-1 $\beta$  基因 +3953C/T 多态性, 根据限制性内切酶 *Taq* I 酶切片段的情况, 观察到基因型有两种: CC 型 (135 bp、114 bp 2 条带), CT 型 (135 bp、114 bp 和 249 bp 3 条带) (见图 1-3)。

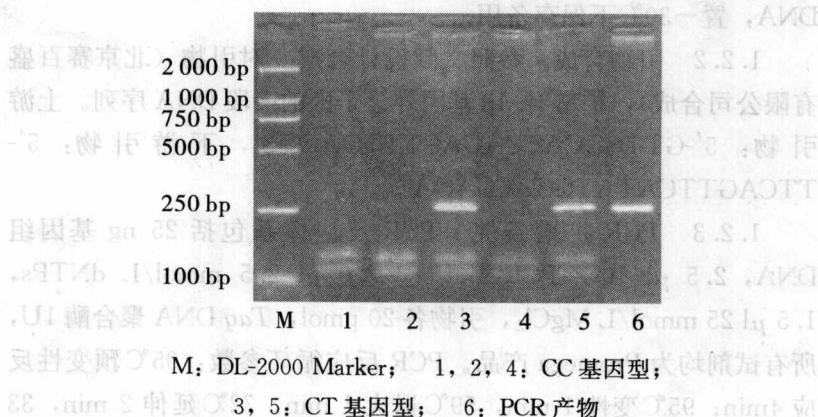


图 1-3 IL-1 $\beta$  (+3953) 基因多态性电泳图