

双链置换探针实时 PCR 用于 DNA 池的 等位基因频率测定

程金平¹, 梁基选², 李庆阁^{1*}

(1. 厦门大学生命科学学院 细胞生物学与肿瘤细胞工程教育部重点实验室,

2. 厦门大学医学院, 抗癌研究中心, 福建 厦门 361005)

摘要: 结合荧光双链置换探针的特异性和实时 PCR 的准确定量能力, 建立了一种新颖、准确、价廉且高通量的测定 DNA 池等位基因频率的方法。该方法采用不同荧光标记的双链置换探针 1 次 PCR 反应即可测定出等位基因频率。实验以 β -地中海贫血 CD_s41-42(-TCTT) 突变为对象, 分别用荧光染料 FAM 和 ROX 标记野生型和突变型探针, 由实时 PCR 检测建立的等位基因浓度与循环阈值的响应曲线计算 DNA 池的等位基因频率。结果表明, 建立的系列等位基因浓度与循环阈值呈线性关系, 线性相关系数分别为 0.997 7(野生型等位基因)和 0.993 8(突变型等位基因), 可检测的最低等位基因频率达 1%, 等位基因频率在 1%~90% 范围内测定误差小于 4%。该方法可广泛用于流行病学调查、遗传连锁分析以及全基因组连锁不平衡扫描等领域。

关键词: 双链置换探针; 实时 PCR; DNA 池; 等位基因频率

中图分类号: Q 52

文献标识码: A

人类基因组计划的完成, 加速了人类致病基因及其突变的鉴定和发现, 不断增长的基因突变信息又为致病基因的大规模筛查、遗传病的关联分析以及大范围的人群的基因诊断奠定了基础。随着各种高通量筛查技术的出现, 常规的针对个体标本的基因分型技术, 无论在效率还是试剂、人力以及 DNA 标本消耗方面都暴露出明显的局限性, DNA 池^[1] 技术为克服这一局限提供了解决途径。对 DNA 池进行基因分型, 可以大大减少实验数目, 节省劳动量、时间和费用, DNA 池已在多种高通量 SNP 基因分型中得以应用^[2]。

应用 DNA 池检测基因突变和测定等位基因频率已涉及到多种突变检测技术, 尤以高通量技术倍

受青睐, 如变性高效液相色谱法(DHPLC)^[3~5]、物质谱法^[6,7]、基因芯片法^[8]、生物发光法^[9]等。然而由于这些方法大都涉及多步后 PCR 操作, 不仅过程繁琐, 且降低了等位基因频率测定的准确性, 应用受到很大限制。实时 PCR 的应用则在很大程度上克服了上述缺点^[10,11]。

荧光双链置换探针具有高度特异性^[12,13], 可以在室温下进行单个核苷酸多态性的基因分型^[14]。实验选择 β -地中海贫血的 CD_s41-42(-TCTT) 突变为研究对象, 考察了荧光双链置换探针实时 PCR 法结合 DNA 池技术用于测定等位基因频率的可行性。

1 材料与方法

1.1 材料

1.1.1 DNA 模板

β -珠蛋白野生型质粒直接以野生型基因组 DNA 为模板构建, CD_s41-42(-TCTT) 突变质粒采用引物重叠延伸法构建。构建的质粒均由上海生工生物工程有限公司双向测序确认。人基因组 DNA 模板由广州第一军医大学徐湘民教授惠赠。

1.1.2 引物与探针

上下游引物分别为: 5'-TCTGCCTATTG-

收稿日期: 2004-03-25

基金项目: 国家自然科学基金资助项目(30170834)和福建省自然科学基金重点项目(C0220001)资助。

作者简介: 程金平(1978-), 女, 硕士研究生。

* Corresponding author, Tel/Fax: 0592-2182100,

E-mail: qqli@xmu.edu.cn

GTC TATTTTCCC -3' 和 5'-CATAACA GCATCAG - GAGTGGACA 3', 扩增 β 珠蛋白基因 CDs41-42 位点附近一段 110 bp 的片断. 野生型和突变型荧光双链置换探针分别为: 5'-FAM-GACCCAGAGGT TCTTTGAGTCCTT-PO₄ 3' / 5'-AGGACTCA AAGA ACCTCTGGGTG-Dabcyl 3' 和 5'-Texasred-GAC-CCAGAGG TT GAGT CCTTTGGG-PO₄ 3' / 5'-CCAAGGACTC AACCTCTGGGTG-Dabcyl 3', 下划线部分表示突变位置. 引物和探针分别由上海生工生物工程有限公司和上海申友生物工程有限公司合成.

1.2 方法

1.2.1 DNA 池的构建

参照文献[2, 10] 方法设计: 质粒 DNA 经紫外分光定量, 稀释至 10^{10} copies/mL. 通过混合不同体积的野生型和突变型质粒模板构建不同等位基因频率的 DNA 池, 构建的野生型和突变型的等位基因频率包括 0%、1%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%.

1.2.2 双链探针实时 PCR 的特异性实验

分别用野生型和突变型人基因组 DNA 作为模板对双链置换探针实时 PCR 检测突变的特异性进行了考察. 所用模板包括 2 份野生型模板(平行管)和 2 份突变型模板(平行管), 水作为阴性控制. 在 25 μ L PCR 反应液中, 含 2.5 μ L 10 \times PCR 缓冲液, 3.0 mmol/L MgCl₂, 200 μ mol/L dNTP, 2.0 U Taq 酶(rTaq, 大连宝生物工程有限公司), 0.4 μ mol/L 引物, 0.2 μ mol/L 野生型和突变型探针(双链置换探针按荧光链与淬灭链 1:1.2 比例混合而成)和 10 μ L DNA 模板(50 ng/ μ L). PCR 反应在实时 PCR 仪(Rotor Gene 3000, Corbett Research)上进行, 首先 95 $^{\circ}$ C 变性 3 min, 然后经 94 $^{\circ}$ C 变性 15 s, 52 $^{\circ}$ C 退火 20 s, 72 $^{\circ}$ C 延伸 25 s, 共 45 个循环, 设置在每一循环的退火阶段用 FAM 和 ROX 频道分别采集野生型和突变型探针的荧光数据.

1.2.3 双链置换探针实时 PCR 检测的线性关系实验

考察了实时 PCR 检测模板浓度的情况, 模板的浓度梯度为: 10^6 、 10^7 、 10^8 、 10^9 、 10^{10} 、 10^{11} 、 10^{12} 、 10^{13} (copies/mL), 水为空白对照. 反应体系和反应条件同上. 实验分别用野生型和突变型两种探针针对两种模板进行了考察, 实时 PCR 检测条件同 1.2.2.

1.2.4 等位基因频率测定

双链置换探针实时 PCR 双频道同时检测了野

生型和突变型的等位基因频率, 模板包括 1.2.1 中构建的系列频率的 DNA 池, 反应体系和反应条件同上. 依据实时 PCR 检测的结果建立定量工作曲线, 得到 C_t 值与等位基因浓度之间的对应关系, 然后计算出等位基因浓度占总 DNA 池的浓度的百分数, 此百分数就是该等位基因的频率. 在测定等位基因频率时, 对上述构建的每个 DNA 池平行测定两次, 取两次检测的平均值为检测结果, 按照上面建立的工作曲线来计算系列 DNA 池的等位基因频率, 实验测定的频率与已知频率之间的差值作为相对误差.

2 结果

2.1 双链置换探针实时 PCR 检测基因突变

实验表明, 荧光双链置换探针能够清楚区分 CDs41-42 的野生型模板和突变型模板(图 1), 且平行管的 C_t 值一致. 显示出荧光双链置换探针的特异性和实时 PCR 的准确性. 平行实验还显示反应平台期之后, 荧光强度存在差异, 这一差异正是所有建立在 PCR 产物分析的方法定量欠佳的原因所在. 实时 PCR 采用反应刚刚进入指数扩增的标志 C_t 作为定量依据, 显然更为准确.

2.2 双链置换探针检测浓度梯度模板

应用 DNA 池必然涉及到不同浓度模板的稀释, 因此要求测定应有一个宽广的定量范围. 实时 PCR 检测结果显示, 在 $10^3 \sim 10^{10}$ (copies/reaction) 的范围内, 模板量与其 C_t 值呈线性关系, 线性相关系数分别为: 野生型 $R^2 = 0.9934$, 突变型 $R^2 = 0.9864$ (R 为线性相关系数)(图 2).

分别用野生型双链置换探针和突变型双链置换探针检测了 9 个系列梯度构建的质粒模板的 PCR 反应, 模板浓度从高到低依次为 10^{10} 、 10^9 、 10^8 、 10^7 、 10^6 、 10^5 、 10^4 、 10^3 , 以及阴性对照. (a) 野生型双链置换探针检测野生型模板实时 PCR 灵敏度实验; (b) 野生型双链置换探针检测野生型模板的标准曲线; (c) 突变型双链置换探针检测突变型模板实时 PCR 灵敏度实验; (d) 突变型双链置换探针检测突变型模板的标准曲线.

2.3 双链置换探针检测突变以及测定等位基因频率

实时 PCR 运用类似梯度模板定量的方法来测定 DNA 池中等位基因的浓度, 然后根据等位基因浓度与总的 DNA 池浓度的比值计算出等位基因频

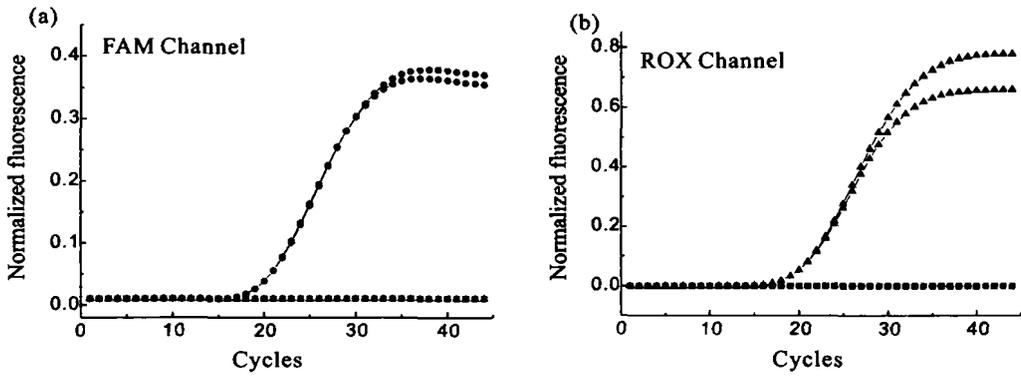


图1 荧光双链置换探针实时 PCR 基因分型

(a)野生型双链置换探针在 FAM 通道产生的荧光信号;(b)突变型双链置换探针在 ROX 通道产生的荧光信号。●-●:野生型模板;▲-▲:突变纯合子模板;■-■:阴性对照。

Fig.1 Real time PCR genotyping with double stranded displacing probes

率,因此等位基因频率的测定实际上就是等位基因浓度的测定.从图 3(a)和 3(c)可以看出,等位基因频率与 Ct 值相对应,从 3(b)和 3(d)可以看出, Ct 值与等位基因浓度相对应,而等位基因的浓度占总浓度的百分数就是该等位基因的频率,如此就可以根据不同频率对应的等位基因的浓度做的定量标准

工作曲线检测 DNA 池标本的等位基因频率.从图 3 还可以看出,获得的定量标准工作曲线呈现良好的线性关系,线性范围覆盖了从 1%到 90%的等位基因频率范围.在具体测定某一个 DNA 池中的野生型等位基因频率时,首先从实时 PCR 的定量曲线可以得到其 Ct 值,然后在定量工作曲线上找到这

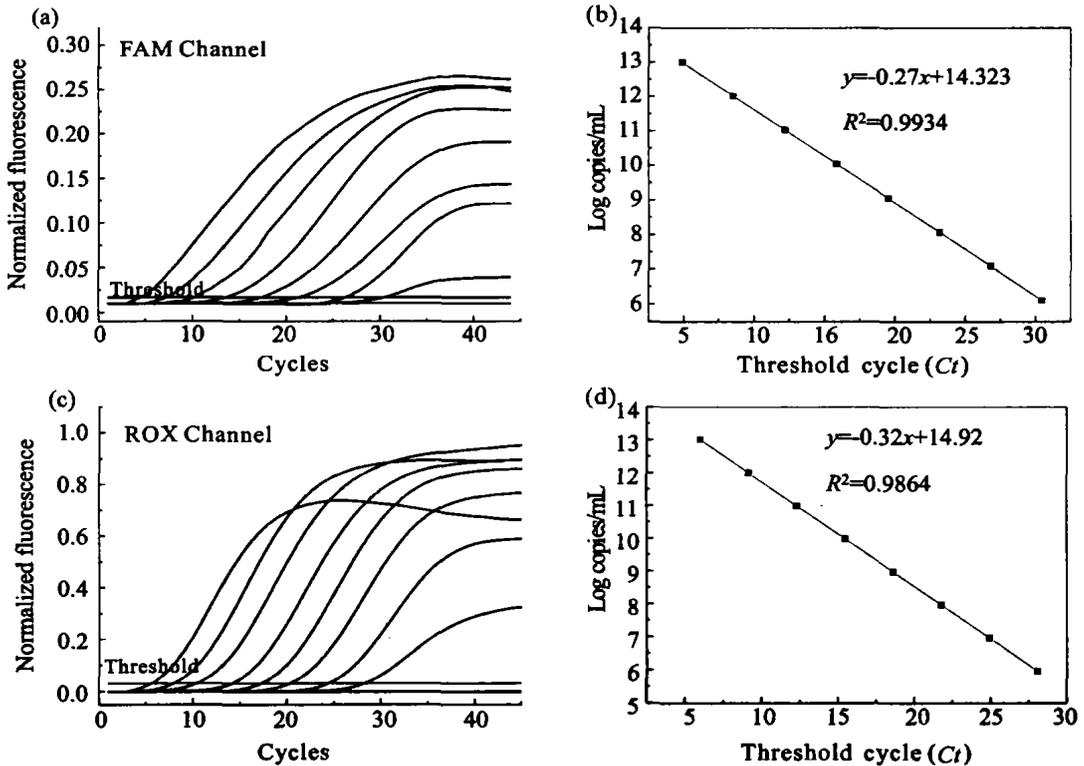


图2 荧光双链置换探针实时 PCR 灵敏度实验

Fig.2 Determination of the Sensitivity of the assay with double stranded displacing probes

个 C_t 值对应的等位基因的浓度,再由已知总的 DNA 池的浓度,就可以计算出其基因频率.一般认为检测到了某一个等位基因的频率也就同时知道了另一个等位基因的频率^[10,11].另外,采用同一反应双频道检测模式,还可以相互印证各自检测的等位基因频率,使结果更加可靠.

用一系列总浓度相同而等位基因频率不等的模板来进行荧光双链置换探针实时 PCR 测定等位基因频率实验.(a)野生型双链置换探针实时 PCR 测定野生型等位基因频率;(b)野生型双链置换探针测定野生型等位基因浓度的标准曲线;(c)突变型双链置换探针实时 PCR 测定突变型等位基因频率;(d)突变型双链置换探针测定突变型等位基因浓度的标准曲线.

2.4 等位基因频率测定的误差分析

采用 2.3 中的方法来计算等位基因频率,针对两个等位基因的 11 个频率测定的相对误差分析表明:野生型等位基因的平均误差为 1.6%,突变型等位基因的平均误差为 1.48%.野生型等位基因的最大误差为 3.83%,最小误差为 0.04%;突变型等位基因的最大误差为 2.84%,最小误差为 0.33%.基

本的规律是 60%~90% 这一个区间的基因频率的检测误差稍大(表 1).总之,双链置换探针 DNA 池技术测定等位基因频率的准确率稍高于 96%,与 ARMS-PCR 的 95%^[10]相近.

3 讨论

荧光双链置换探针实时 PCR 技术能够可靠地检测点突变,并且可以在宽广的 DNA 浓度范围内实现准确的定量检测,结合样品池技术最低能检测 100 个标本构建的样品池中 1% 的基因突变,达到甚至超过了目前 DNA 池技术基因频率检测的灵敏度.荧光双链置换探针具有的设计简便、价格便宜的优点,也特别适合 DNA 池的高通量的要求.

目前用于 DNA 池的突变检测技术原理上包括限制性酶切片断多态性(RFLP),引物特异延伸,探针杂交以及单链构象多态性(SSCP)^[2].RFLP 因为酶切不完全的问题,SSCP 因为准确性低的缺点,都已很少用于 DNA 池.基于引物延伸原理的 DHPLC 和物质谱技术具有通量高且分型准确的优点,测定等位基因频率时,DHPLC 技术利用峰形的高低

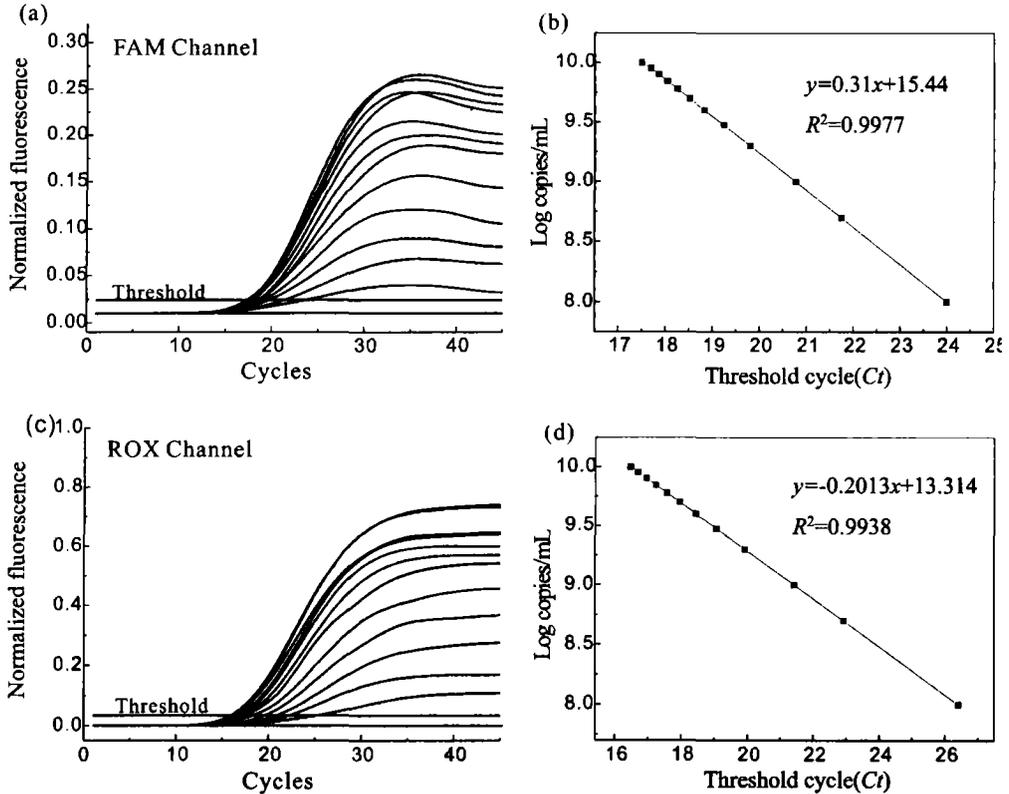


图 3 荧光双链置换探针实时 PCR 测定等位基因频率实验

Fig. 3 Allele frequencies determination with double stranded displacing probe based real time PCR using a series of templates of the same DNA concentration but different relative allele amount

表 1 双链置换探针实时 PCR 测定不同频率的等位基因频率实验结果以及误差分析

Tab. 1 Allele frequency determination with double stranded displacing probes across a range of values

已知的等位基因 频率(OD_{260})/%	预期 C_t 值		测定 C_t 值 ^a		测定的等位基因频率/%		平均误差(±)/%	
	WA ^b	MA ^c	WA	MA	WA	MA	WA	MA
1	24.00	26.40	24.06	26.49	0.96	0.51	0.04	0.49
5	21.75	22.93	21.70	22.85	5.19	4.67	0.19	0.33
10	20.77	21.43	20.82	21.51	9.65	10.61	0.35	0.61
20	19.80	19.94	19.75	19.85	20.83	19.42	0.83	0.58
30	19.24	19.06	19.29	19.15	28.78	28.86	1.22	1.14
40	18.83	18.44	18.78	18.36	41.59	41.45	1.59	1.45
50	18.52	17.96	18.47	17.88	51.88	51.33	1.88	1.33
60	18.26	17.57	18.31	17.64	57.98	57.27	2.02	2.73
70	18.05	17.23	18.10	17.32	67.37	72.56	2.63	2.56
80	17.86	16.94	17.81	16.86	83.05	82.19	3.05	2.19
90	17.70	16.69	17.65	16.60	93.83	92.84	3.83	2.84

测定 C_t 值^a: 取 2 个平行管测定 C_t 值的平均值; WA^b: 野生型等位基因; MA^c: 突变型等位基因。

来测定等位基因频率的大小, 生物质谱则须先建立等浓度的不同碱基的分子量标准, 然后根据检测峰的面积求和计算出等位基因频率. 这两种技术的共同缺点是步骤繁琐, 定量信号易受仪器条件干扰, 尤其是最后的分析步骤是建立在 PCR 产物基础上, 所得出的等位基因频率并不直接代表 DNA 池中模板的频率. 同样地, 建立在探针杂交基础上的微阵列技术用于 DNA 池也存在类似问题. 上述问题使这些高通量方法在 DNA 池中的应用受到很大限制.

采用实时 PCR 模式, 无论是依据引物延伸原理还是探针杂交原理, 都具有操作简便和定量准确的优点. 相对而言, 基于引物延伸原理的 ARMS-PCR 兼有特异扩增功能更容易拉大不同等位基因之间的差别. 但目前采用荧光染料的方式引入了非特异扩增和两次反应的问题^[10]. 双色荧光探针技术使这些问题迎刃而解, 但 DNA 池技术对探针的特异性要求很高, 目前仅有的报道是采用 TaqMan MGB 探针^[11]. TaqMan MGB 探针的缺点是由于修饰繁多(三重修饰)而价格昂贵, 目前尚难以普及. 因此, 我们提出采用荧光双链置换探针以克服这一弊端, 结果表明, 后者完全可以满足 DNA 池技术的要求. 另外, 选择多重染料进行标记, 还可以做到多个 SNPs 的同时检测. 值得指出的是, 探针的特异检测技术并不具备特异扩增的优势, 因此, 很难进一步提高少数等位基因频率检测的灵敏度, 但对于常规 DNA 池来说, 目前的灵敏度已经足够. 可以预期, 采用多色荧光修饰的等位基因特异引物将可既克服 ARMS-

PCR 的缺点, 又保留其特异扩增优势, 有望进一步提高基因频率检测的灵敏度.

参考文献:

- [1] 童大跃. 混合 DNA 样品池扩增法及其应用[J]. 生物技术通讯, 1999, 10(3): 213-216.
- [2] Sham P, Bader J S, Craig I, et al. DNA Pooling: a tool for large scale association studies [J]. Nat. Rev. Genet., 2002, 3: 862-871.
- [3] Wolford J K, Blunt D, Ballecer C, et al. High throughput SNP detection by using DNA pooling and denaturing high performance liquid chromatography (DHPLC) [J]. Hum. Genet., 2000, 107: 483-487.
- [4] Giordano M, Mellai M, Hoogendoorn B, et al. Determination of SNP allele frequencies in pooled DNAs by primer extension genotyping and denaturing high performance liquid chromatography [J]. J Biochem Biophys Methods, 2001, 47: 101-110.
- [5] Hoogendoorn B, Norton N, Kirov G, et al. Cheap, accurate and rapid allele frequency estimation of single nucleotide polymorphisms by primer extension and DHPLC in DNA pools [J]. Hum. Genet., 2000, 107: 488-493.
- [6] Werner M, Sych M, Herbon N, et al. Large scale determination of SNP allele frequencies in DNA pools using MALDI TOF mass spectrometry [J]. Hum. Mutat., 2002, 20(1): 57-64.
- [7] Mohlke K L, Erdos M R, Scott L J, et al. High throughput screening for evidence of association by using mass spectrometry genotyping on DNA pools [J]. Proc. Natl. A-

- cad. Sci. USA, 2002, 99(26): 16 928– 16 933.
- [8] Uhl G, Liu Q R, Walther W, et al. Polysubstance abuse vulnerability genes: genome scans for association, using 1 004 subjects and 1 494 single nucleotide polymorphisms [J]. Am. J. Hum. Genet., 2001, 69: 1 290– 1 300.
- [9] Zhou G, Kamahori M, Okano K, et al. Quantitative detection of single nucleotide polymorphisms for a pooled sample by a bioluminometric assay coupled with modified primer extension reactions (BAMPER) [J]. Nucleic Acids Res., 2001, 29: E93.
- [10] Germer S, Holland M J, Higuchi R. High throughput SNP allele frequency determination in pooled DNA samples by kinetic PCR [J]. Genome Res., 2000, 10 (2): 258– 266.
- [11] Xu K, Lipsky R H, Mangal W, et al. Single nucleotide polymorphism allele frequencies determined by quantitative kinetic assay of pooled DNA [J]. Clin. Chem., 2002, 48: 1 605– 1 608.
- [12] Li Q G, Luan G Y, Guo Q P, et al. A new class of homogeneous nucleic acid probes based on specific displacement hybridization [J]. Nucleic Acids Res., 2002, 30(2): E5.
- [13] Cheng J, Zhang Y, Li Q. Real time PCR genotyping using displacing probes [J]. Nucleic Acids Res., 2004, 32(7): E61.
- [14] 张永有, 李庆阁, 栾国彦, 等. 双链探针同步荧光技术快速筛查 C282Y 点突变 [J]. 遗传, 2003, 25(1): 9– 13.

Determination of Allele Frequency of Pooled DNA by Double-stranded Displacing Probe-based Real Time PCR

CHENG Jir ping¹, LIANG Jir xuan², LI Qing ge^{1*}

(1. The Key Laboratory of the Education Ministry of Cell Biology and Tumor Cell Engineering, School of Life Sciences, Xiamen University, 2. Cancer Research Center, Medical College, Xiamen University, Xiamen 361005, China)

Abstract: An accurate, inexpensive and high throughput method for determining the allele frequency of biallelic polymorphisms in pools of DNA samples was developed. The assay combined the high specificity of double-stranded displacing probes with real-time quantitative PCR. Using differently labeled probes, the relative amounts of each allele in a sample could be quantified in one reaction. This method was evaluated using CD41-42(-TCTT) mutation in human β -globin gene causative for β -thalassemia as an example. The wild-type and mutant probes were labeled with FAM and ROX, respectively, and were recruited in one reaction. Using the constructed DNA pools, real time PCR was performed to establish the relationship between the relative amount of each allele and the threshold cycle value. With this relationship, allele frequency of the pooled DNA sample could be obtained by its threshold cycle value and the DNA concentration of the sample. The results showed that there is a linear relationship between the amount of the allele and the threshold cycle value through the whole allele frequency range studied, and the lowest allele frequency detected was 1%. The linear correlation coefficient was 0.997 7 for wild-type allele and 0.993 8 for mutant allele, respectively. Within the allele frequency range of 1% ~ 90%, the relative error calculated was less than 4%, which was acceptable for allele frequency determination in pooled DNA samples. Considering its simplicity, reliability, and low cost, this approach was a powerful strategy for large population-based epidemiology study, detecting meaningful polymorphic differences in candidate gene association studies, and genome wide linkage disequilibrium scans.

Key words: double-stranded displacing probe; real-time PCR; DNA pooling; allele frequency