

学校编码: 10384

分类号_____密级_____

学号: 200326081

UDC _____

廈門大學

硕士学位论文

探针编码实时 PCR 的基因分型研究

Development of Real-time PCR Genotyping Methods Based on
Probe Coding Technology

指导教师姓名:

专业名称: 细胞生物学

论文提交日期:

论文答辩时间:

学位授予日期:

答辩委员会主席: _____

评 阅 人: _____

2006 年 月

厦门大学学位论文原创性声明

兹提交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其他个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文产生的权利和责任。

声明人（签名）：

年 月 日

厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

- 1、保密（ ），在 年解密后适用本授权书。
- 2、不保密（ ）。

（请在以上相应括号内打“√”）

作者签名： 日期： 年 月 日

导师签名： 日期： 年 月 日

目 录

中文摘要	1
英文摘要	3
第一章 探针编码技术基因分型概述	6
第一节 基因分型的意义和现有分型技术的局限性	6
第二节 探针编码技术基因分型原理	7
§ 2.1 新型荧光双链置换探针实时 PCR 基因分型的原理	7
§ 2.2 探针编码技术的基本原理	8
§ 2.3 探针编码技术的设计规则	10
参考文献	11
第二章 基于探针编码技术的蜡样芽孢杆菌分型研究	15
第一节 引言	15
第一节 材料和方法	16
第三节 结果	18
§ 3.1 扩增产物的验证	18
§ 3.2 反应体系的改进	18
§ 3.3 蜡样芽孢杆菌分型体系的建立	20
§ 3.4 7 株蜡样芽孢杆菌分型结果	21
第四节 讨论	22
参考文献	23
第三章 基于探针编码技术的 HPV 分型研究	26
第一节 引言	26
第一节 材料和方法	28
第三节 结果	35
§ 3.1 15 种 HPV 基因型质粒的验证	35
§ 3.2 荧光双链置换探针实时 PCR 的特异性实验	36
§ 3.3 2 份临床标本的检测	39
§ 3.4 双重感染模拟实验	39
第四节 讨论	41
参考文献	43

第四章 探针编码技术结合假病毒技术用于 HCV 基因实时

分型.....	50
第一节 引言	50
第一节 材料和方法	53
第三节 结果	58
§ 3.1 6 种 HCV 基因型假病毒表达载体 PCR 验证和测序.....	58
§ 3.2 假病毒的包裹实验.....	58
§ 3.3 实时 RT-PCR 探针编码基因分型体系的建立.....	59
§ 3.4 临床样品的检测.....	60
第四节 讨论	60
参考文献	61
硕士期间发表交流论文	68
致谢.....	69

CONTENTS

Abstract(In Chinese)	1
Abstract(In English)	3
Chapter I The summary of Probe coding technology	6
Section I Introduction	6
Section II Principle of PCT based genotyping	7
§2.1 Principle of real-time PCR genotyping using displacing probes	7
§2.2 Basic principle of PCT	8
§2.3 The design formula of PCT	10
References	11
Chapter II Bacillus cereus genotyping based on PCT	15
Section I Introduction	15
Section II Materials and Methods	16
Section III Results	18
§3.1 PCR amplification results of 15 Bacillus cereus strains	18
§3.2 Optimization of the genotyping assay.....	18
§3.3 Real-time PCR genotyping of 8 standard strains of Bacillus cereus with PCT	20
§3.2 Genotyping results of 7 samples of known genotypes.....	21
Section IV Discussion	22
References	23
Chapter III HPV genotyping based on PCT	26
Section I Introduction	26
Section II Materials and Methods	28
Section III Results	35
§3.1 Amplification products for 15 HPV genotypes.....	35

CONTENTS

§3.2 Specificity of displacing probes in real-time PCR genotyping	36
§3.3 Real-time PCR genotyping of 2 clinical samples.....	39
§3.4 Double-infection simulate test	39
Section IV Discussion	41
References	43
Chapter IV HCV genotyping based on VLP technology and PCT	50
Section I Introduction	50
Section II Materials and Methods	53
Section III Results	58
§3.1 PCR results of 6 VLP vector and sequencing analysis.....	58
§3.2 The package of VLP	58
§3.3 Construction of the real-time genotyping system	59
§3.4 RT-PCR genotyping of 10 clinical samples	60
Section IV Discussion	60
References	61
Publications	68
Acknowledgment	79

摘要

本论文围绕探针编码技术基因分型展开了研究，研究工作主要包括探针编码技术的提出和基本原理，以及探针编码技术在三种不同研究对象基因分型中的应用。

第一章，提出了一种用于基因分型的新型探针编码技术，并对其基本原理和设计规则进行了系统的说明。探针编码技术通过不同荧光染料的组合标记靶序列，利用实时 PCR 过程中采集到的荧光信号来判断基因型。它可以超越实时 PCR 仪器检测通道的限制，具有高通量，简单快速，特异性强等优点，极大地提高了实时 PCR 基因分型的能力。

第二章，以食源性致病菌蜡样芽孢杆菌为研究对象，建立了探针编码技术实时 PCR 基因分型的检测体系，并进一步应用于 7 株已知基因型的菌液标本的分型，均得到与测序一致的结果。

第三章，应用探针编码技术区分人乳头瘤病毒 15 种基因型。系统地考察了探针编码技术在四色荧光 PCR 仪器上进行基因分型所能达到的理论极限情况，即用 4 种荧光染料区分 15 种基因型，并在实验中得以实现。另外，还模拟临床中可能出现的 HPV 双重感染情况，通过调整两种染料荧光值的比例，对其进行了准确的区分。在此基础上，完成两份临床标本的基因分型。

第四章，以 RNA 假病毒作为质控品结合探针编码技术，对中国地区常见的 6 种基因型进行区分，建立了简单快捷，便于质控的 HCV RT-PCR 分型体系，并对 10 份厦门地区 ELISA 检测阳性的临床标本进行分型，均可准确确定基因型。

与传统的基因分型技术相比，探针编码技术具有设计简单，特异性强，

高通量等优点。探针编码技术的检测通量接近于低通量基因芯片，但操作和结果判断更为简便，可以广泛应用于基因分型领域。

关键词：置换探针；探针编码；基因分型

厦门大学博硕士学位论文摘要库

ABSTRACT

This dissertation consists of four chapters. The first chapter describes the development and principle of probe coding technology (PCT), the following three chapters deal with the application of PCT genotyping in *Bacillus cereus*, human papillomavirus, and hepatitis C virus.

In chapter one, a novel real-time PCR genotyping approach named probe coding technology (PCT) was described. The basic principle of PCT is that each genotype to be analyzed is coded by from one to four fluorophores-labeled displacing probes, which could discriminate sequences difference at single nucleotide level. In such a way, using n fluorophores, totally 2^n-1 targets can be coded. The genotyping results can be obtained directly after real-time PCR using these probes based on the appearance of fluorescence corresponding to the probes. PCT can be regarded as a multiplex real-time PCR genotyping strategy concerning the targets number to be analyzed. PCT actually combines the merits of high-throughout and simplicity of real-time PCR, the extremely high specificity of displacement probes, and the legible mathematics rule, and has thus greatly improved the genotyping capacity of real-time PCR.

In chapter two, the feasibility of PCT was tested by detecting eight different genotypes of a food-born pathogen *Bacillus cereus*. A four-step protocol and asymmetric PCR was adopted to facilitate the hybridization between the probes and targets. The results showed that using totally nine probes labeled with four different colors, i.e., FAM, HEX, ROX and Cy5, the eight genotypes *Bacillus cereus* could be unequivocally discriminated within 2

hr. The robustness of PCT was further validated by genotyping of seven strains isolated from food samples and the results were confirmed with bi-directional sequencing.

In chapter three, a multiple real-time PCR genotyping of fifteen human papillomavirus in a single-tube was developed. The design reached the theoretical upper limit of PCT genotyping, i.e., 15 genotypes were discriminated using 4 fluorophores-labeled probes. The results showed that using thirty-two probes labeled with different combination of FAM, HEX, ROX, and Cy5, 15 HPV genotypes could be discriminated within 2 hr. In addition, differentiation of co-infection was also tested by changing the ratio the fluorescence of two probes. The established method was applied to two clinical samples. The results showed that one sample was HPV6, the other was HPV16. These results were concordant with the sequencing method.

In the last chapter, PCT was applied to reverse-transcription PCR genotyping of HCV. First, 6 armored RNA of HCV fragments were prepared standing for the 6 different genotypes mostly popular among Chinese population. Using 9 probes labeled with three different colors, i.e., HEX, ROX and Cy5, the 6 HCV genotypes could be discriminated within 3 hr. Finally, 10 clinical samples were genotyped and confirmed by sequencing.

In comparison with conventional genotyping methods, PCT is more accurate, reliable, and high-throughput. The analytical capacity of PCT is close to low density microarray or chip of the same aim, but PCT has obvious advantages in manipulation and readout interpretation. It is more suitable for routine clinical testing and large-scale genetic screening.

Keywords: Displacing Probes; Probe Coding Technology; Real-time PCR Genotyping.

厦门大学博硕士学位论文摘要库

第一章 探针编码技术基因分型概述

第一节 基因分型的意义和现有分型技术的局限性

基因多态性, 又称DNA多态性, 是指不同物种或同一物种不同个体间基因组序列的差异, 包括基因编码区和非编码区序列的差异。它是生物多样性的遗传基础, 决定了物种之间和物种内部个体间的差异。例如, 同一种微生物可以根据基因组差异的大小分为多个亚型, 人类基因组单核苷酸多态性 (SNPs) 也使个体呈现多样表型。从分子诊断的角度看, 对病原微生物进行基因分型, 可以准确溯源, 协助医生对患者实施针对性治疗; 对人类基因进行SNPs分型, 可以进行个体识别、遗传病筛查, 并对器官移植以及个性化治疗加以指导^[1-3]。

由于基因分型研究具有重要的意义, 各种分型技术应运而生^[4,5]。这些方法主要有限制性片段长度多态性 (RFLP)^[6], 单链构象多态性 (SSCP) 分析技术^[7], 基于DNA多态性产生异源双链的变性梯度凝胶电泳 (DGGE)^[8], 异源双链分析 (HA)^[9]和变性高效液相色谱 (DHPLC)^[10]等分析技术, 基于探针杂交的等位基因特异性寡核苷酸分析法 (ASO)^[11], RDB反向点杂交^[12]和生物芯片^[13]等分析技术, 基于引物特异扩增或特异延伸的ARMS-PCR^[14], 突变引物延伸扩增 (MOEA)^[15]和MALDI-TOF-MS^[16]分析技术, 基于测序的Sanger测序^[17], Minisequencing^[18]和Pyrosequencing^[19]序列分析技术, 以及近年来发展的MLPA和MIP探针技术^[20,21]。虽然检测通量较大, 但是大部分的分析技术都需要进行PCR后处理, 分析相对比较繁琐, 而且PCR后处理操作容易带来污染, 因而制约了大部分技术的发展以及在临床诊断领域的应用。另外, 精细的基因分型依赖于对基因序列差异程度的准确识别, 随着基因型数目的不断积累, 用于识别的位点数也随之增加。因此, 发展能够在单核苷酸水平上识别多个基因位点的基因分型技术, 就成为分子诊断研究的重要任务。

作为无需PCR后处理的最重要的均相检测技术, 实时PCR在基因定量、基因分型和SNPs检测等方面正得到越来越广泛的应用^[22]。实时PCR一个突出优点是扩增检测同步进行, 其它核酸扩增依赖的检测技术, 无论检测的通量和

操作如何,其简便程度都不及实时PCR,这也是目前实时PCR在临床上得以普遍应用的根本原因。然而,实时PCR存在检测容量小的问题,即受仪器可检测的荧光种类数目的限制,难以在一次反应中同时检测多个靶序列。最近上市的实时PCR仪器能同时检测5种不同荧光染料,按目前的检测方式至多能检测5种靶序列。这一问题严重制约了实时PCR在基因分型中的应用。

实时PCR的简便性及其诸多优点吸引人们不断努力去提高其检测容量。自Kramer小组^[23]率先在四色荧光PCR仪器实现HTLV1、HTLV2、HIV-1和HIV-2四种靶序列的同时检测以来,多组分实时PCR检测技术已广泛应用。2000年Tyagi等^[24]发明波长迁移型分子信标,克服了单一光源无法同时检测多个荧光染料的困难,El Hajj等^[25]据此实现了5色探针检测结核杆菌的利福平耐药突变。

撇开实时PCR仪器, Lee等^[26]利用荧光光度计的同步扫描7种荧光染料,实现了6种靶序列的同时检测。此外,利用PCR扩增产物的熔点差异,在实时PCR扩增结束后,进行熔点曲线分析,也可实现多个靶序列的同时检测和基因分型^[27,28]。后者的局限性在于PCR产物的熔点范围狭窄,很难容下5个以上产物的同时存在,而且容易受到引物二聚体等非特异扩增的干扰。迄今,单个反应能进行5种基因型以上分型的实时PCR尚未实现。

针对多重实时核酸扩增检测容量小,即在一次反应中可同时检测的靶序列数目受仪器荧光检测通道数目的限制的问题,我们拟建立一种使目前的多重实时核酸扩增可检测的靶序列数目明显多于目前仪器可检测的荧光染料数目的探针编码技术(Probe Coding Technology, PCT)。

第二节 探针编码技术基因分型原理

§ 2.1 新型荧光双链置换探针实时PCR基因分型的原理

实时PCR应用于DNA多态性基因分型的关键技术是探针技术。目前基于实时PCR的各种标记探针技术主要包括TaqMan探针^[29], FRET探针^[30], 分子信标探针^[31]及新型荧光双链置换探针^[32,33]。FRET相邻探针需要在PCR扩增完成后进行熔点分析进行基因分型,并非实时判断。而普通的TaqMan探针很难达到

基因分型要求的特异性，于是引进了MGB以增强特异性，但是也增加了设计合成的难度以及成本的费用。分子信标应用于基因分型方面也需要复杂的设计，而且其在单碱基识别方面存在缺陷。因此这几种探针无法很好的满足实时PCR基因分型的需要。

新型荧光双链置换探针是两条按照脱氧核糖核酸（DNA）碱基互补配对原则而反向互补的寡核苷酸。在其中的一条寡核苷酸的5'端标记荧光供体（也可称荧光基团），在另一条寡核苷酸的5'端标记荧光受体（也可称淬灭基团），同时两条寡核苷酸的3'端进行封闭（如磷酸化），使其不能作为引物延伸。根据荧光能量转移原理，当新型荧光双链置换探针在不与靶序列发生置换杂交时，荧光供体与荧光受体很靠近，二者可以发生荧光能量传递，荧光基团的荧光被淬灭，此时检测不到荧光；当新型荧光双链置换探针标记荧光供体的链和标记荧光受体的链分别与靶序列发生置换杂交时，荧光供体与荧光受体由于新型荧光双链置换探针两条链的分离而分开，荧光基团的荧光无法被淬灭基团淬灭，此时就可以检测到荧光信号。因此通过荧光信号的变化就可以知道荧光双链置换探针两条寡核苷酸链所处的状态。由于荧光双链置换探针自身两条链之间的互补配对，存在着与靶序列互补的竞争作用，因此荧光双链置换探针具有非常好的识别单碱基突变的能力，而且其特异性可以通过调节探针的长度和两条链相差的碱基数来实现，具有很强的特异性和灵敏度，因此适用于基因分型。

由于标记不同荧光染料，与相应基因型完全匹配的双链置换探针与特定基因型目的片段杂交可以发出特定的荧光信号，通过实时聚合酶链式反应过程中产生的荧光信号来判断检测标本的基因型，从而达到基因分型的目的。

荧光双链置换探针的高特异性和灵敏度，为探针编码技术的提出和应用奠定了坚实的理论基础。

§ 2.2 探针编码技术的基本原理

探针编码技术（Probe Coding Technology, PCT）的基本原理，简言之，就是对于任一基因型的特异探针，通过标记不同种类的荧光元素，使其获得一个独特编码，将所有不同编码的探针混入同一反应管，通过实时PCR给出的编码探针信号进行基因分型。

编码是将荧光基本元素组合成不同的元素构成，比如，A 元素和 B 元素，可以分别构成 A、B 和 A+B；同样 A、B 和 C 元素，可以构成 A、B、C、A+B、A+C、B+C、A+B+C 共七种组合，这种组合方式符合数学上的组合规则。因此，按照数学上的组合规则，对于 N 个荧光元素，共可以得到 $C_N^1 + C_N^2 + \dots + C_N^N = \sum C_N^i (i=1 \sim N) = 2^N - 1$ 种组合方式，其中的任何一种组合方式都可以标记一种靶序列特异探针，故可得到 $2^N - 1$ 种不同标记的荧光探针，这样就可以检测 $2^N - 1$ 种不同的靶序列，并以二进制进行编码。按这一方案，四色荧光 PCR 仪器（共有四种荧光基本元素）就可检测 $2^4 - 1 = 15$ 种靶序列，五色荧光 PCR 仪器（共有五种荧光基本元素）就可检测 $2^5 - 1 = 31$ 种基因型。显然，按该规则，当前的实时 PCR 检测模式只占组合编码的其中一项即 C_N^1 。

以标记四种荧光染料的情况为例，探针编码技术基因分型原理如图 1 所示。

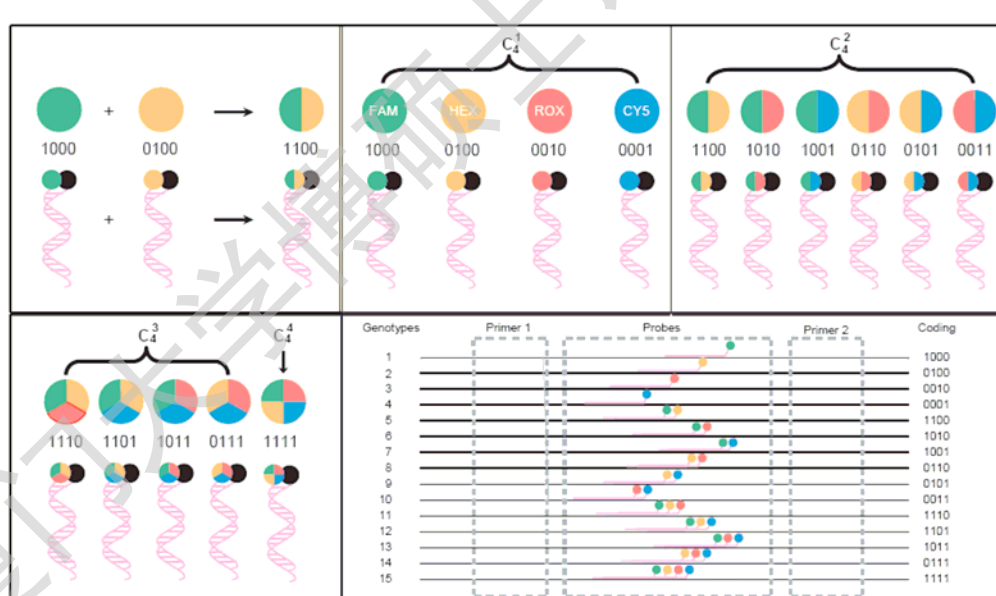


Figure 1. Schematic representation of PCT genotyping strategy. There are four differently colored fluorophores: FAM (in green), HEX (in Orange), ROX (in red) and CY5 (in blue).

从图中可知，四种荧光染料分别为 FAM，HEX，ROX 和 CY5。当进行单色标记时，可以区分四种基因型；如果两种染料共同标记一种基因型，就可以组合出 FAM+HEX，FAM+ROX，FAM+CY5，HEX+ROX，HEX+CY5 和 ROX+CY5 这六种

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士学位论文摘要库