

文章编号: 0529-6579 (1999) 05-0061-06

HLA DQ 位点多样性模拟分析及其 PCR-RFLP 分型^{*}

陈尚武¹, 冯炳健¹, 潘德京¹, 陈荣祥², 徐安龙¹

(1. 中山大学生命科学学院, 广州 510275;

2. 云南澜沧县第一人民医院)

摘 要: 选择 27 个识别序列为 4 nt 的限制性内切酶, 通过计算机对最新获得的 19 个 DQA1 和 35 个 DQB1 等位基因第 2 外显子中的一段序列进行模拟酶切分析, 根据酶切格局数目及各格局所代表的等位基因数的均衡性, 确定酶的优先次序, 在此基础上进行 PCR-RFLP 模拟分析, 确定最佳的酶组合. 结果表明, 分别采用 4 个和 6 个酶的组合, 可有效鉴定 DQA1 和 DQB1 等位基因. 同时通过比较分析 DQB1 各等位基因核苷酸位点的变异系数, 发现 DQB1 不同血清型间等位基因第 2 外显子核苷酸序列变化有一定规律. 研究为建立一个简捷准确的 HLA 分型方法提供了新的理论基础.

关键词: HLA; 基因型; 计算机分析

中图分类号: R 392 **文献标识码:** A

HLA (human leukocyte antigen) 是人类最富多样性的遗传系统, 广泛应用于人类学研究 and 法医个体识别, 同时参与抗原提呈和同种异体移植物的识别. HLA 多样性与人类疾病相关. HLA 抗原的分型由血清学和细胞学的方法发展到抗原蛋白质电泳、单克隆抗体分子、RFLP、直接测序和以聚合酶链反应 (PCR) 技术为基础的一系列方法.

HLA DQ 抗原编码 α 链 (DQA) 和 β 链 (DQB), 其第二外显子与各种抗原结合, 所以该位点在人群中存在一定的多态性. 某些特定的等位基因与一些疾病具有相关性^[1~3], 因而研究 HLA DQ 位点的基因分型具有重要意义. 1989 年, Maeda^[4] 首先建立了 PCR-RFLP HLA 基因分型方法: PCR 扩增出 HLA 某一位点, 然后用一组限制性内切酶消化, 根据核苷酸片段大小, 电泳分析而定型. 随着 HLA 各位点所发现的等位基因的数目越来越多, 酶切格局越来越复杂. 旧的 PCR-RFLP 只能鉴定较少的等位基因, 使用的内切酶也较多, 并常有较多等位基因无法区分. 根据最新 DQ 位点序列资料^[6,7], 对 19 个 DQA1 和 35 个 DQB1 PCR-RFLP 基因分型的计算机模拟, 并对 DQB1 各血清型间等位基因核苷酸序列的变化规律进行了初步分析, 为研究 HLA 结构与功能的关系提供了依据.

^{*} 基金项目: 国家杰出青年科学基金 (39725007) 资助项目; 广东省自然科学基金 (980366) 资助项目

收稿日期: 1998-11-27 作者简介: 男, 1965 年生, 副教授 通讯联系人: 徐安龙

1 材料与方法

1.1 DQA1 和 DQB1 等位基因序列资料来源

以 IHWC 网址下载的截止 1998 年 3 月共 19 个 DQA1 和 35 个 DQB1 等位基因为分析对象, 所有等位基因不再进行分组。

1.2 引物序列

DQA1 为 DQAF2: TCAGCTGACCATGTTGCCTCT, DQAR1: GGACACATACCATTGGTGGC, 扩增第 2 外显子, 长度为 260 或 263 bp. DQB1 为 GH28: CTCGGATCCGCATGTGCTACTTCACCAACG, GH29: GAGCTGCAGGTAGTGTGTCTGCACAC, 扩增第 2 外显子, 长度为 230 bp.

1.3 分析软件及思路

(1) 选取 27 个识别序列为 4 碱基的 II 类限制性内切酶, 分别对 DQA1 和 DQB1 等位基因进行模拟酶切分析, 根据不同酶所产生的切点数、酶切格局数及酶切格局分布的均衡性计算出各酶的优先系数, 排出各酶的优先次序。

(2) 以酶组合可鉴定的等位基因最多为原则, 分别选取 4 个和 6 个内切酶依次对 DQA1 和 DQB1 等位基因进行 PCR-RFLP 模拟分析, 逐步计算出可能鉴定的等位基因, 编制相应的酶切格局表, 根据酶切格局确定相应的等位基因。对一些仍不能定型或区分的等位基因, 选取另外一些识别序列为 5~8 nt 的限制性内切酶进行分析。

(3) 区别 DQB1 不同血清型的等位基因, 分别计算扩增产物长度内每一组等位基因各碱基的变异系数(待发表, 变异系数值在 0~1 之间, 0 表示最保守, 即该位点在同组各等位基因间没有差别; 1 表示该位点 A, T, G, C 4 个碱基在各等位基因间分布是均匀的), 绘制出碱基变异系数图。

2 结果与讨论

2.1 27 种内切酶对 19 个 DQA1 和 35 个 DQB1 等位基因的酶切结果

表 1 中, 酶切格局指该酶将等位基因区分的组数目; 表观方差为不同酶切格局所代表等位基因数目的方差, 转换成 0~1 后, 取 1 与其差值; 优先指数为表观方差值与酶切格局数目的乘积, 决定酶在表中的位置; 效率指按表中顺序每加入新的酶后的酶组合, 累计可鉴定的等位基因百分数; 效果指按表中顺序每加入新的酶可将全部等位基因再次区分为新的组别数; 酶识别序列中, J 代表 G 或 C 碱基; H 代表 A 或 T 碱基; N 代表 A、T、C、G 4 种碱基中的任意 1 种。

表观方差: 设等位基因数目为 M , 酶切格局数目为 N , 酶切格局中每组数目为 X_i ,
表观方差 = $1 - \{ \sum (X_i - M/N)^2 / [(M - M/N)^2 + (N - 1)(0 - M/N)^2] \}^{1/2}$

当只有 1 种酶切格局时, 方差最大。代入上述公式时, 分母为 0, 无意义, 故直接定义其为 1。

结果分析表明, DQB1 (结果从略) 等位基因的多态性比 DQA1 高, 其中的酶切点数与酶切格局数目比 DQA1 复杂。优先指数表明该酶有较多的酶切格局数目, 且每一种格局所代表的等位基因数目比较均匀, 单独使用应该优选, 但由于各酶间在酶切格局所代表的等位基因数目上可能存在重叠, 故酶组合的选择重点应考虑酶组合所能鉴定的等位基因数目。

表 1 27 种内切酶对 DQA1 等位基因 PCR 产物 (260 bp) 的酶切结果

Tah 1 Twenty-weven enzyme cleaving result of DQB1's PCR product

内切酶	切点数	酶切格局数	表观方差	优先指数	效率	效果
RsaI (GTAC)	2	2	0.947	1.893	0	1
Tsp45I (GTJAC)	1	2	0.947	1.893	0.2	2
AciI (CCGC)	2	2	0.593	1.186	0.3	1
MnlI (CCTC)	3	3	0.593	1.778	0.4	1
TaiI (ACGT)	1	2	0.787	1.573	0.4	0
Sau96I (GGNCC)	2	2	0.787	1.573	0.4	0
MspI (CGG)	1	2	0.787	1.573	0.4	0
Tsp509I (AATT)	3	2	0.787	1.573	0.4	0
AvaII (GGHCC)	2	2	0.787	1.573	0.4	0
BsaI (CCNNGG)	1	2	0.787	1.573	0.4	0
DdeI (CTNAG)	2	2	0.787	1.573	0.4	0
NlaIV(GGNCC)	1	2	0.787	1.573	0.4	0
SerFI (CCNGG)	1	1	1	1	0.4	0
BstNI (CCHGG)	1	1	1	1	0.4	0
Fnu4HI (GCNGC)	1	1	1	1	0.4	0
HhaI (GCGC)	0	1	1	1	0.4	0
HaeIII(GGCC)	1	1	1	1	0.4	0
AluI (AGCT)	1	1	1	1	0.4	0
BfaI (CTAG)	0	1	1	1	0.4	0
BstUI (CGCG)	0	1	1	1	0.4	0
MseI (TTAA)	1	1	1	1	0.4	0
NlaIII(CATG)	2	1	1	1	0.4	0
TaqI (TGA)	0	1	1	1	0.4	0
Cac8I (GCNNGC)	0	1	1	1	0.4	0
HirFI (GANTC)	0	1	1	1	0.4	0
NciI (CCJGG)	0	1	1	1	0.4	0
TscI (GCHGC)	1	1	1	1	0.4	0

2.2 内切酶分别对 19 个 DQA1 和 35 个 DQB1 等位基因的 PCR-RFLP 模拟分析

应用一组 4 种内切酶,可鉴定出 7 种 DQA1 等位基因(表 2).其中,DQA1 *06011 和 *06012 核酸序列不同,但它们编码的蛋白序列相同,视为一种不作区分.

表 2 4 种酶对 19 个DQA1 等位基因的酶切格局

Tah 2 Four enzymes cleaving pattem of 19 DQA1 alleles

Rsa I	Tsp45 I	Aci I	MnlI	等位基因
37 223	260			* 0201
	69 191			* 06011 * 06012
27 37 196	69 191			* 0401
	260	23 73 164		* 0501 * 0503 * 0504
		23 237		* 0502
27 37 199	263	27 28 73 135	126 44 23 17 62	* 0101
			126 44 40 62	* 0102
		22 23 73 145		* 03011 * 0302 * 0303
37 226				* 0103

表 3 6 种酶对 35 个DQB1 等位基因的酶切格局
Tab 3 Six enzymes cleaving pattern of 35 DQB1 alleles

Hha I	Nla VI	ScrFI	Aci I	Fnu4HI	Tsp45I	等位基因
39 48 143	6 48 176	11 45 60 114				*0501
		11 45 174				*05031
	6 224	56 174				*05032
		56 60 114	4, 120, 15, 57, 3, 31			*0201 *0202
			4, 120, 72, 3, 31			*0203
	6 47 177	56 174	4, 119, 1 17, 55, 34	5 9 90 126 5 90 135		*06011 *06013 *06012
39 48 56 87	6 48 176					*0502
	6 224					*0504
31 39 46 114	6 48 176	11 45 174	4, 108 13, 1, 17, 55 34	5 9 90 126	52 69 109 109 121	*0602 *0611 *0613 *0603 *0608 *0614
	6 47 177					*0607
	39 46 145	6 47 177	11 45 60 114	4, 119, 1 8, 64 34	5 9 90 126	
5 9 35 55 126					*0606	
6 48 176					*0612	
6 47 50 127					*03012	
6 26 47 50 101					*0305	
6 26 198					*0402	
27 31 39 46 87						*0610
39 191	6 47 50 127	11 45 60 114	4, 119, 1 8, 64, 34			*03011
			4, 119, 1 8, 7 57, 34			*0304
		11 45 174	4, 119, 1 8, 7 57, 34			*0302 *0307
			4, 119, 1 8, 64, 34			*03032
		6 97 127				
85 145						*0401

但DQA1 *01 中的 *0101, *0104, *0105; DQA1 *05 中的 *0501, *0503 以及 DQA1 *03 (*0301, *0302, 0303)彼此不能区分. 选用 6 种酶可鉴定出 19 个 DQB1 等位基因(表 3), 另有 2 对等位基因和 3 个 3 等位基因组(蛋白质序列相同未计算在内)彼此不能区分. 对这些彼此不能区分的等位基因, 可选用其他的内切酶进行分析, 如使用 BseRI 可区分 DQA1 中的 *0102 与 *0101, *0104, *0105. 计算机模拟分析结果表明, 选用适当的内切酶, 通过 PCR-RFLP 可有效区分 HLA DQ 位点的等位基因.

PCR-RFLP 用于 HLA 基因分型已有应用, 但使用的酶比较多且比较特殊^[6~8]. 我们的结果与以往的 PCR-RFLP 相比, 使用的酶没有明显的增加(Mizuki 等 DQA1 用了 6 种酶, DQB1 用了 10 种酶, 且其中大部分种类稀少, 价格不菲), 但通过对酶的优化分析和选用, 可区分的等位基因数增加了. 更为重要的是, 由于发现的等位基因的数量不断增加, 通过计算机分析, 根据现有的等位基因资料, 建立一个比较规范的 PCR-RFLP HLA 基因分型方法十分必要.

3 血清学分型与等位基因核苷酸序列的关系

比较 DQB1 不同血清型等位基因核苷酸序列的变异系数图(图 1)可以发现, 大多数等位

基因序列都在比较固定的位置出现变化,但图的右边(序列靠近第2 内含子),在各血清型内是保守的,而在不同的血清型间却有较大的变异,表明这可能是各血清型之间抗原性之差异所在,也可能与特定的等位基因个体对疾病的易感性或抗性不同有关,同时也是根据血清型设计分组引物的优选部位.

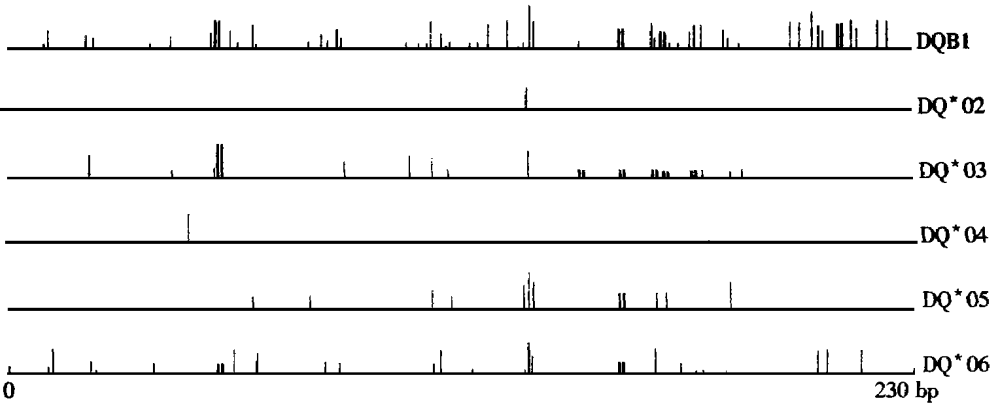


图 1 DQB1 不同血清型等位基因核苷酸序列的改变
Fig 1 Difference index of different serotype of DQB1

图中每一栏中的横线从左到右为扩增的 DQB1 第二外显子序列,竖线长度代表碱基变异的幅度,第一栏为 35 个 DQB1 等位基因序列总的变化情况,2~6 栏为 DQB1 *02~*06 血清型等位基列的变异情况.变异系数=1- 各种碱基数目的方差/ 各种碱基数目方差的最大值.

目前,除 PCR-RFLP,基于 PCR 的 HLA 基因分型方法还有 PCR-SSP (PCR-sequence specific primers)、PCR-SSCP (PCR-single strand conformation polymorphism)、PCR-SSO (PCR-sequence specific oligonucleotide)等几种,但一般都需要很多特异的引物或探针,需要严格的 PCR 条件或杂交条件,基层单位难以推广.而且,根据现有的等位基因序列设计的引物和探针,不能检出新的等位基因.从 PCR-RFLP 新带型的出现,可以发现新的等位基因.PCR-RFLP 以其简便、快捷、廉价、不需同位素等优点成为 HLA 基因分型中倍受重视的方法之一.

参考文献:

[1] 彭学标,岳晓玉,徐文严,等. HLA-DQ 多态性基因与系统性红斑狼疮产易感性[J] . 上海免疫学杂志, 1997, 17(4): 241.

[2] 张修武,张之炯,郭迅,等. 原发性肥厚型心肌病与 HLA-DQ 的相关性研究[J] . 上海免疫学杂志, 1997, 17(1): 22.

[3] 张薇,陈楠,董德长,等. HLA-DQA1, DQB1 基因与特发性膜性肾病关联研究[J] . 中华微生物学与免疫学杂志, 1996, 16(6): 435.

[4] MAEDA M, MURAGAMA N, ISHII H, et al. A simple and method for HLA-DQA1 genotyping by digestion of PCR-amplified DNA with allele specific restriction endonucleases[J] . Tissue Antigens, 1989, 34(5): 290.

[5] MIZUKI N, OHNO S, ANDO H, et al. Major histocompatibility complex class(II) [J] . Tissue Antigens, 1998, 51 (3): 287.

[6] 张咸宁,帕孜来提,邵红光,等. 用 PCR-RFLP 方法研究新疆地区汉族 HLA-DQA1 和-DQB1 基因多态性

- [J]. 中国免疫学杂志, 1996, 12(2): 89.
- [7] 陆建荣, 潘华星, 仁远, 等. 用 PCR-RFLP 方法研究浙江沪籍汉族 HLA-DQA1 基因多态性[J]. 中国免疫学杂志, 1994, 10(6): 367.
- [8] 黄啸宇, 王亚新, 王福庆, 等. HLA-DRB1 基因位点多态性的 PCR-RFLP 分析[J]. 生物化学杂志, 1996, 12(6): 674.

Computer Analysis of Human HLA DQ Locus Polymorphism and Genotyping by PCR-RFLP

CHEN Shang-wu^{*}, FENG Bing-jian, PAN De-jing, CHEN Rong-xiang, XU An-long

Abstract: The exon 2 sequences of 19 DQA1 and 35 DQB1 alleles bounded respectively on two pairs of primers were cleaved successively with a set of 27 restriction endonucleotidases with 4nt recognition sequence by using a new computer program in order to determine the enzyme cleaving pattern and select proper enzymes for genotyping. The genotyping of DQA1 and DQB1 alleles by polymerase chain reaction—restriction fragment length polymorphism (PCR-RFLP) was also performed with computer program. Results demonstrate a set of 4 and 6 restriction endonuclease have efficiently genotyped DQA1 and DQB1 alleles by PCR-RFLP, respectively. Some alleles can be identified with other restriction endonuclease besides above 27 enzymes. The analysis of polymorphism of DQB1 alleles suggest that variability of nucleotide sequences of exon II among DQB1 alleles of different serotypes presents some rules. The study provide the theoretical base for DQB1 and DQA1 genotyping by PCR-RFLP.

Keywords: HLA; genotype; computer analysis

^{*} School of Life Sciences Zhongshan University, Guangzhou 510275, China