

中国汉族人群人类白细胞抗原 HLA 不同基因座位间的基因重组

骆 媛[†] 袁 方[†] 孙玉英^{†*} 金 荔 梁 飞 刘 楠 宋新强
刘曙光 刘金锋 王 丹 薛 红 孔繁华 奚永志^{*}

(军事医学科学院附属医院免疫学实验室及国家生物医学分析中心 免疫学实验室, 北京 100071)

摘要 人类白细胞抗原(human leucocyte antigen, HLA)复合体基因座位间的重组、重组热点、连锁不平衡与疾病易感性的关联性研究已成为免疫遗传学与人类后基因组学等研究领域的研究重点与焦点. 本文报道了在中国汉族人群中发现的 10 例 HLA 复合体中的基因重组家系. 采集重组家系的 2~3 代家庭成员的外周血标本, 首先对其 HLA-I, II 类区域的 5 个经典基因位点进行 PCR-SSP 低、高分辨率基因分型及 SBT 分析, 然后再进行遗传家系分析研究, 确定 HLA 基因重组相关位点, 并对重组家系的单体型进行了群体统计学分析. 结果发现, 在 10 例 HLA 重组家系中有 8 例重组发生在 HLA-A/Cw 位点之间, 其中 4 例重组发生在 A*02-Cw*03 单体型中, 2 例重组发生在 A*24-Cw*03 单体型中; 1 例重组发生在 HLA-B/DRB1 位点之间, 还有 1 例重组发生在 HLA-B/Cw 位点之间. 群体统计学分析发现, A*02-Cw*03 单体型呈一定程度的连锁不平衡但不是很强($\Delta/\delta_{\Delta} = 3.0$, $\Delta_{(r)} = 0.30$); 另有 3 个单体型虽呈强连锁不平衡却发生了重组事件, 即 A*33-Cw*03 ($\Delta/\delta_{\Delta} = 7.22$, $\Delta_{(r)} = 0.51$), A*30-Cw*06 ($\Delta/\delta_{\Delta} = 10.37$, $\Delta_{(r)} = 0.93$)和 B*13-Cw*03 ($\Delta/\delta_{\Delta} = 7.69$, $\Delta_{(r)} = 0.38$), 这说明重组可能是打破旧的连锁不平衡来产生新的连锁不平衡的重要原因. 本文绘制了每个重组家系成员的 HLA 单体型, 并建立了中国汉族人群的 HLA 重组遗传图谱, 初步定位重组发生的范围并分析了位点间重组与连锁不平衡的关系, 这为更深入地研究 HLA 重组事件奠定了基础.

关键词 人类白细胞抗原(HLA) 基因重组 家系研究 连锁不平衡

人类主要组织相容性复合体(major histocompatibility complex, MHC), 又称人类白细胞抗原(human leucocyte antigens, HLA)复合体, 是人类基因组内基因密度较高的一个区域, 位于人类第 6 号染色体短臂 6p21.31, 基因片段长约 4 Mb, 占人体整个基因组的 1/3000. 这些基因在机体防御外界细菌和病毒入侵、器官移植配型以及与自体免疫和某些传染性疾病相

关联方面具有十分重要的意义 [1,2]. MHC 因其基因突变、链间交换以及染色体重组(交换)等机制使其成为人体最复杂的遗传多态性系统, 截至 2006 年 6 月, HLA I 及 II 类等位基因已达到 2550 个. MHC 多态性是人体面对纷繁多样致病微生物所形成的一种防御性保护机制, 在人类抵御外来微生物侵袭、保护机体免受损伤等方面具有重要的意义. 因此, 近 20 年来

收稿日期: 2006-12-10; 接受日期: 2007-03-21

国家重大基础研究发展计划(批准号: 2003CB515509)、国家自然科学基金(批准号: 3047051)及军事医学科学院“创新”基金重大专项资助项目

[†] 同等贡献

* 联系人, E-mail: xiyz@yahoo.com, yuyings@yahoo.com

对MHC重组热点以及重组、连锁不平衡与疾病易感性的关联性研究就成为免疫遗传学与人类后基因组学等研究领域研究重点。尤其是近几年来,大量相关研究论文均发表在*Nature*及其系列与*Science*等著名杂志上^[3-5]。由此,不同种族MHC遗传多态性与重组热点相关研究的重要性是显而易见的。此外,2003年10月英国科学家成功破译人类第6号染色体遗传密码的巨大进展,更加完善了人们对于人类MHC的深入研究与认识^[6]。

业已证明,人类主要组织相容性复合体(MHC)与100多种疾病相关。基础及临床研究资料均表明,疾病与MHC的关联具有人群或种族特异性,而连锁不平衡(LD)分析是研究疾病与遗传标志之间是否关联的最有力的遗传学工具,通过LD分析可定位致病基因。应特别指出的是,基因重组在群体不同LD单倍型的形成过程中起到了关键性作用,所以只有深入研究已有的和不断发现的重组事件,才能真正揭示LD形成的机制与分布规律。另外大量研究资料表明,HLA重组不是随机发生的,而且具有种族特异性,某些种族特异性的HLA单倍型或基因多态性与其种族易感染某些疾病有很强的关联性。尤为重要的是,我国人口占世界人口的1/5,民族众多,从多种自身免疫性疾病如强直性脊柱炎、I型糖尿病以及传染性疾病如艾滋病、SARS以及禽流感等疾病的发病特点以及MHC的关联来看^[7],我国汉族人群与世界其他人群具有显著的不同。因此,进行中国汉族人群MHC重组事件的研究对于我国人类遗传学研究、常见及重大疾病的关联与防治研究具有重大的基础理论及临床应用价值。本研究以10例HLA重组家系为起点,初步研究了我国汉族人群人类白细胞抗原复合体中的基因重组事件以及重组与连锁不平衡之间的关系。

1 材料和方法

1.1 研究对象

收集自2000年以来在我室进行造血干细胞移植前HLA配型的供受体外周血标本5000例共计1002个家系,其中,HLA复合体重组家系10例及其2~3代家庭成员;并随机抽取1014例汉族人群标本进行群体遗传学分析。除重组家系成员外,所有标本均已进行HLA-A、-B、-Cw、-DRB1及-DQB1基因座位的

低分辨率基因分型。

1.2 研究方法

(1) DNA的提取: 采用QIAGEN公司提供的QIAamp DNA Blood Mini kit提取。取0.5%EDTA抗凝全血200 μ L至1.5 mL离心管,加入200 μ L AL缓冲液和20 μ L蛋白酶K混匀,56℃孵育10 min,加入200 μ L无水乙醇混匀,转移至树脂柱,8000 r/min离心1 min,加入500 μ L AW1,8000 r/min离心1 min,加入500 μ L AW2,13000 r/min离心3 min。将树脂柱转移至1.5 mL离心管中,加入200 μ L无菌水,室温孵育1 min,8000 r/min离心1 min。紫外分光光度计(BECKMAN)测 A_{260}/A_{280} 为1.7~1.9,DNA定量为50~100 ng/ μ L。

(2) HLA-I, II类PCR-SSP低分辨率及高分辨率基因分型^[8]: 分别采用美国戴诺公司提供的PEL-FREEZ ABC SSP™(Lot#026)及DRDQ 3T SSP™(Lot#002)分型试剂盒对10个重组家系的成员进行HLA-A、B、Cw、DRB1及DQB1位点的低分辨率分型;高分辨分型亦采用该公司提供的各个位点的高分辨率分型试剂盒。按照试剂盒说明将适量DNA、蒸馏水、Taq酶与PCR缓冲液混合均匀后,取8 μ L混合液加入分型板的反应孔中,阴性对照孔不加,用密封条封闭试剂板,并标记样品号。将试剂板放在PE-9700 PCR仪上进行PCR反应,反应参数为(96℃, 60 s), 1个循环→(96℃, 25 s; 70℃, 50 s; 72℃, 45 s), 5个循环→(96℃, 25 s; 65℃, 50 s; 72℃, 45 s), 21个循环→(96℃, 25 s; 55℃, 60 s; 72℃, 120 s), 4个循环。取PCR反应产物8 μ L上样到2%的琼脂糖凝胶,150 V电压下电泳4 min。在紫外凝胶成像仪下观察结果,分析电泳带的格局,确定等位基因。

(3) SBT(sequence based typing)分型: HLA-B/Cw位点重组家系中,由于目前HLA-B位点高分辨试剂盒不能有效区分HLA-B*1301及B*1302,故采用SBT法对B位点全长基因组进行测序分型,并比较父、女B位点基因组差异,以判断重组是否发生在B基因座位上。用HLA-B位点通用引物^[9]5'-GGAC-TCAGAGTCTCCTCAGACGCCGAG-3'和5'-CTGGG-GAGGAAACACAGCTCAGCATGGGAAC-3'扩增B位点全长基因组。PCR扩增体系为: 10×缓冲液5 μ L, dNTP (2.5 mmol/L) 4 μ L, gDNA 3 μ L, 引物各1 μ L, Pyrobest (TaKaRa)高保真酶(5 U/ μ L) 0.25 μ L, 无菌水

35.75 μL , 总体积 50 μL . 循环参数为: (94 $^{\circ}\text{C}$, 30 s; 66 $^{\circ}\text{C}$, 30 s; 72 $^{\circ}\text{C}$, 3 min), 10 个循环 $\rightarrow$ (94 $^{\circ}\text{C}$, 30 s; 64 $^{\circ}\text{C}$, 30 s; 72 $^{\circ}\text{C}$, 3 min), 20 个循环 $\rightarrow$ 72 $^{\circ}\text{C}$, 10 min. 用 DNA 回收纯化试剂盒(Marligen)回收纯化 PCR 产物后, 将其插入 pGEM-T easy (Promega) 载体, 转化大肠杆菌 DH5 α , 选取阳性克隆送上海博亚公司进行测序.

(4) 群体遗传学统计分析方法: 从本研究室收集的拟进行造血干细胞移植的 5000 例供-受体样本中随机选择 1014 例汉族个体的 HLA 等位基因资料, 计算两位点单体型频率. 其估算公式如下:

$$s = 1 + \sqrt{d/N} - \sqrt{B/N} - \sqrt{D/N},$$

连锁不平衡参数 $\Delta = s - p_i p_j$, Δ 的标准误

$$\delta_{(A)} = \sqrt{\frac{a}{4N^2} - \frac{\Delta}{N} \left(\frac{B+D}{2\sqrt{BD}} - \frac{\sqrt{BD}}{N} \right)},$$

相对连锁不平衡参数 $A_{(r)} = \Delta / A_{(\text{最大})}$, 在 $\Delta / \delta_{(A)} \geq 1.96$ 时, 认为是显著连锁不平衡^[10]. 此数据用来估算重组单体型连锁不平衡程度, 从而验证重组的发生是否与单体型连锁不平衡程度有关.

2 结果

2.1 患者及其家庭成员的 HLA 分型结果

在获得 HLA 低分辨率基因分型结果的基础上, 为进一步确定这些家系的 HLA 等位基因, 我们又进行了 HLA 高分辨率基因分型实验, 结果详见表 1.

2.2 HLA-B 位点全长基因组测序结果

由于在对第 10 个家系 HLA-B 位点进行高分辨率分型时, 不能将 B*1301 及 B*1302 有效分离, 故对该家系 3 个成员的 HLA-B 基因座位进行全基因组克隆测序, 由此能够将 PCR-SSP 高分辨方法难以分辨的 B*1301 和 B*1302 清楚的区分开来, 保证了分型结果的准确性. 同时将父、女的测序结果进行对比, 发现 2 人的 B*1301 座位基因组序列并无差异, 提示重组不是发生在 B 位点上, 而是出现在 B/Cw 座位之间的区域.

2.3 重组家系的单体型遗传分析研究

为进一步确证家系重组的发生, 我们先进行 HLA 单体型定型, 然后再采用单体型遗传分析进行研究. 我们把父亲的 2 条单倍体设为 a 和 b, 母亲的设为 c 和 d, 重组后的单倍体为 a/b 或 c/d. 减数分裂

的过程中, 来自父源或母源的染色单体在遗传给子代时, 等位基因的线性排列组合方式却发生了变化, 在子代的染色单体中均找不到与父亲或母亲一致的单体型, 分析是由于重组而造成. 如表 1 所示, 家系 1~8, 均是父亲或母亲的 2 条染色单体在其减数分裂的过程中 HLA-A/Cw 位点之间发生了交换, 重组后产生了新的单体型 a/b 或 c/d. 家系 1 中, 一子女的母源单体上的 A*330301/0302 来自母亲的一条染色体, 而 C*030401/0402-B*4601-DRB1*090102/02-DQB1*030302/0303 则来自母亲的另外一条染色体, 这种新的线性排列组合与母亲/父亲的单倍体型别都不同, 即为母亲染色体重组的结果. 家系 2 中, 父亲的 2 条单倍体亦在 HLA-A/Cw 位点之间发生了重组, 重组后所产生新的单体型 A*24020101-0205-C*07020101/020101-B*15010101/020102-DRB1*0901-05-DQB1*030302/0303 遗传给了女儿. 与家系 1, 2 类似, 家系 3~8 都是父亲或母亲的 2 条染色单体在减数分裂的过程中在 HLA-A/Cw 位点间发生了重组, 重组所产生的新单体型遗传给了子女. 家系 9 中, 由于父母在国外, 血样采集困难, 故双亲的单倍体型是按照家系单体型遗传分析原理由全部子女的单体型推断出来的. 其中一个子女的单体型 A*330301/0302-C*010201/0202-B*4601-DRB1*12021-022-DQB1*030101/0102 与所有同胞的 a, b, c, d 4 个单体型均不同, 通过家系单体型遗传分析排除是外来基因的可能, 实际上是父亲/母亲染色体交换的结果. 通过对这个重组先证者女儿的进一步追踪验证, 发现这个新的重组单体型又完整地遗传给了下一代. 家系 10 中, 女儿父源染色单体上的 A*02011, C*150201/0202 来自父亲的一条染色单体, 而 B*1301, DRB1*0406, DQB1*030302/0303 来自父亲的另一条染色单体, 因此推测父亲的 2 条染色单体在减数分裂时 HLA-B 和 -Cw 基因座位间发生了重组, 即 2 条染色体间发生了交换.

2.4 群体遗传统计学对重组单体型连锁不平衡分析

我们对发生重组的 16 种 2 座位单体型通过群体统计学分析(1014 个随机个体)估算其连锁不平衡度. 结果如表 2 所示. 当 $\Delta / \delta_{(A)} \geq 1.96$ 时, 认为是显著连锁不平衡. 16 个重组单体型中, 有 3 个单体型呈强连锁不平衡, 即 A*33-Cw*03($\Delta / \delta_{(A)} = 7.22$, $A_{(r)} = 0.51$)、A*30-Cw*06($\Delta / \delta_{(A)} = 10.37$, $A_{(r)} = 0.93$)和 B*13-Cw*

表 1 10 个重组家系成员每条单体型的 HLA 等位基因型别

家系	单体型	A 位点等位基因	Cw 位点等位基因	B 位点等位基因	DRB1 位点等位基因	DQB1 位点等位基因
1	a	300101/0102	0602	1302	070101/0102	0202
	b	110101/0102	140201/0202	510101-0105	040501-0504	0401
	c	020301/0302	030401/0402	4601	090102/02	030302/0303
	d	330301/0302	030201/0202	5801	030101/0102	020101/0102
	e	020301/0302	030401/0402	4601	090102/02	030302/0303
	a	300101/0102	0602	1302	070101/0102	0202
	d	330301/0302	030201/0202	5801	030101/0102	020101/0102
	b	110101/0102	140201/0202	510101-0105	040501-0504	0401
	c/d	330301/0302	030401/0402	4601	090102/02	030302/0303
	b	110101/0102	140201/0202	510101-0105	040501-0504	0401
2	a	24020101-0205	030301-0303	15010101/0102-0105	130101-0103	0603
	b	02010101-0111	07020101/020102	15010101/0102-0105	0901-05	030302
	c	310102	0706	440301/0302	130101-0103	0603
	d					
	a	24020101-0205	030301-0303	15010101/0102-0105	130101-0103	0603
	c	310102	0706	440301/0302	130101-0103	0603
	a/b	24020101-0205	07020101/020102	15010101/020102	0901-05	030302/0303
	c	310102	0706	440301/0302	130101-0103	0603
	b	02010101-0111	07020101/020102	15010101/0102-0105	0901-05	030302
	c	310102	0706	440301/0302	130101-0103	0603
	b	02010101-0111	07020101/020102	15010101/0102-0105	0901-05	030302
	c	310102	0706	440301/0302	130101-0103	0603
3	a	03010101/0102/0103	0501	44020101-0203	01	050101
	b	110101/0102	010201/0202	5502	15011/012	0602
	c	020101-0109	030401/0402	4002	11011-013	030101/0102
	d	24020101-0205	030301-0303	3501	100101/0102	050101
	d	24020101-0205	030301-0303	350101/0102	100101/0102	050101
	a	03010101/0102/0103	0501	44020101-0203	01	050101
	b	110101/0102	010201/0202	5502	15011/012	0602
	d	24020101-0205	030301-0303	3501	100101/0102	050101
	b	110101/0102	010201/0202	5502	15011/012	0602
	c	020101-0109	030401/0402	4002	11011-013	030101/0102
	b	110101/0102	010201/0202	5502	15011/012	0602
	c	020101-0109	030401/0402	4002	11011-013	030101/0102
	b	110101/0102	010201/0202	5502	15011/012	0602
	c	020101-0109	030401/0402	4002	11011-013	030101/0102
	c/d	24020101-0205	030401/0402	4002	11011-013	030101/0102
	a	030101/0102/0103	0501	44020101-0203	01	050101
4	a	110101	07020101/020102	3905	1401	050301
	b	310102	150201/0202	510201/0202	1403	030101/0102
	c	02010101-0109	030301-0303	15010101/0102-0104	15011-013	050301
	d	020101-0109	030301-0303	350101/0102	15011-013	0602
	a	110101	07020101/020102	3905	1401	050301

续表 1

家系	单体型	A 位点等位基因	Cw 位点等位基因	B 位点等位基因	DRB1 位点等位基因	DQB1 位点等位基因
4	d	020101-0109	030301-0303	350101/0102	15011-013	0602
	a	110101	07020101/020102	3905	1401	050301
	d	020101-0109	030301-0303	350101/0102	15011-013	0602
	a	110101	07020101/020102	3905	1401	050301
	c	02010101-0109	030301-0303	15010101/0102-0104	15011-013	050301
	a/b	310102	07020101/020101	3905	1401	050301
	d	020101-0109	030301-0303	350101/0102	15011-013	0602
5	a	310102	150201/0202	510201/0202	0406	030201/0202
	b	0207	010201/0202	4601	090102-03	030302/0303
	c	3001	0602	1302	070101/0102	0202
	d	310102	030401/0402	1301	12021-022	030101/0102
	c	3001	0602	1302	070101/0102	0202
	b	0207	010201/0202	4601	090102-03	030302/0303
	c	3001	0602	1302	070101/0102	0202
	a	310102	150201/0202	510201/0202	0406	030201/0202
	c	3001	0602	1302	070101/0102	0202
	a	310102	150201/0202	510201/0202	0406	030201/0202
	a	310102	150201/0202	510201/0202	0406	030201/0202
	d	310102	030401/0402	1301	120201/0202	030101/0102
	a	310102	150201/0202	510201/0202	0406	030201/0202
	d	310102	030401/0402	1301	12021-022	030101/0102
	a	310102	150201/0202	510201/0202	0406	020101/0102
	c/d	310102	0602	1302	070101/0102	0202
6	a	03010101/0102/0103	120301/0302	5801	070101/0102	0201
	b	020301/0302	030401/0402	40011	11011-013	030101/0102
	c	330301/0302	0302	3801	030101/0102	0202
	d	0207	0804	40011	0901-05	030302/0303
	a	03010101/0102/0103	120301/0302	5801	070101/0102	0201
	c	330301/0302	0302	3801	030101/0102	0202
	a	03010101/0102/0103	120301/0302	5801	070101/0102	0201
	c	330301/0302	0302	3801	030101/0102	0202
	d	0207	0804	40011	0901-05	030302/0303
	b	020301/0302	030401/0402	40011	11011-013	030101/0102
	a/b	0203	120301/0302	5801	070101/0102	0201
	c	330301/0302	0302	3801	030101/0102	0202
7	a	02010101-0109	030301-0303	151101/1102	0901-05	030302/0303
	b	110101/0102	04010101-0102	1501101/012-014	15011-013	0602
	c	110101/0102	140201/0202	510101-0103/0105	1201	030101/0102
	d	0203	030401/0402	400101-0104	1401	050201/0202
	c	110101/0102	140201/0202	510101-0103/0105	1201	030101/0102
	b	110101/0102	04010101-0102	1501101/012-014	15011-013	0602
	c	110101/0102	140201/0202	510101-0103/0105	1201	030101/0102
	b	110101/0102	04010101-0102	1501101/012-014	15011-013	0602
	a	02010101-0109	030301-0303	151101/1102	0901-05	030302/0303
	d	0203	030401/0402	400101-0104	1401	050201/0202
	a	02010101-0109	030301-0303	151101/1102	0901-05	030302/0303
	d	0203	030401/0402	400101-0104	1401	050201/0202
	a	02010101-0109	030301-0303	151101/1102	0901-05	030302/0303

续表 1

家系	单体型	A 位点等位基因	Cw 位点等位基因	B 位点等位基因	DRB1 位点等位基因	DQB1 位点等位基因
7	c	110101/0102	140201/0202	510101-0103/0105	1201	030101/0102
	a	02010101-0109	030301-0303	151101/1102	0901-05	030302/0303
	c	110101/0102	140201/0202	510101-0103/0105	1201	030101/0102
	c/d	0203	140201	510101-0103/0105	1201	030101/0102
	e	24020101-0205	0202	4601	0901-05	030302/0303
8	a	110101/0102	030401-0403	1301	15011-013	060101-0103
	b	1102	030301-0304	15010101/0102-0105	15011-013	030302/0303
	c	010101/0102	120301-0303	3801	130101-0103	0603
	d	020601-0603	140201/0202	510101-0103/0105	15011-013	030101/0102
	a	110101/0102	030401-0403	1301	15011-013	060101-0103
	d	020601-0603	140201/0202	510101-0103/0105	15011-013	030101/0102
	a	110101/0102	030401-0403	1301	15011-013	060101-0103
	d	020601-0603	140201/0202	510101-0103/0105	15011-013	030101/0102
	a	110101/0102	030401-0403	1301	15011-013	060101-0103
	d	020601-0603	140201/0202	510101-0103/0105	15011-013	030101/0102
	a	110101/0102	030401-0403	1301	15011-013	060101-0103
	c	010101/0102	120301-0303	3801	130101-0103	0603
	a	110101/0102	030401-0403	1301	15011-013	060101-0103
	c	010101/0102	120301-0303	3801	130101-0103	0603
	a	110101/0102	030401-0403	1301	15011-013	060101-0103
	c/d	020601-0603	120301-0303	3801	130101-0103	0603
	d	020601-0603	140201/0202	510101-0103/0105	15011-013	030101/0102
	b	1102	030301-0304	15010101/0102-0105	15011-013	030302/0303
9	a	330301/0302	010201/0202	4601	15021-023	050101
	b	2407	04010101-0102	3505	12021-022	030101/0102
	c	110101/0102	080101/0102	4601	0406	030201/0202
	d	020301/0302	07020101/020102	380201/0202	16021	050201/0202
	c	110101/0102	080101/0102	4601	0406	030201/0202
	a	330301/0302	010201/0202	4601	15021-023	050101
	d	020301/0302	07020101/020102	380201/0202	16021	050201/0202
	a/b	330301/0302	010201/0202	4601	12021-022	030101/0102
	d	0203	07020101/020102	380201/0202	16021	050201/0202
	b	2407	04010101-0102	3505	12021-022	030101/0102
	c	110101/0102	080101/0102	4601	0406	030201/0202
	a	330301/0302	010201/0202	4601	15021-023	050101
	d	0203	07020101/020102	380201/0202	16021	050201/0202
	a	330301/0302	010201/0202	4601	15021-023	050101
10	a	2402101	030401/0402	1301	0406	030302/0303
	b	02011	150201/0202	4002	1405	05031
	c	2406	0602	1302	070101/0102	0202
	d	330301/0302	030201/0202	5801	030101/0102	0202
	a/b	02010101-0111	150201/0202	1301	0406	030302/0303
	c	2406	0602	1302	070101/0102	0202

表 2 重组单体型频率(两座位)及连锁不平衡参数

单倍型	单倍型频率	连锁不平衡参数 Δ	标准误 $\delta_{(\Delta)}$	$\Delta/\delta_{(\Delta)}$	相对连锁不平衡参数 $\Delta_{(r)}$
A*02-Cw*03	0.0546	-0.0239	0.0080	3.0003	0.3044
A*24-Cw*03	0.0376	0.0009	0.0054	0.1709	0.0080
A*33-Cw*03	0.0471	0.0292	0.0040	7.2207	0.5159
A*02-Cw*07	0.0691	0.0112	0.0065	1.7325	0.0937
A*11-Cw*07	0.0463	0.0050	0.0057	0.8651	0.0364
A*31-Cw*15	0.0044	0.0030	0.0016	1.9106	0.0901
A*30-Cw*06	0.0478	0.0434	0.0042	10.3701	0.9298
A*31-Cw*03	0.0090	0.0005	0.0027	0.1993	0.0203
A*03-Cw*12	0.0097	0.0082	0.0021	3.8250	0.2938
A*11-Cw*14	0.0088	-0.0022	0.0031	0.7275	0.2012
A*01-Cw*12	0.0028	0.0009	0.0015	0.6151	0.0257
A*02-Cw*14	0.0157	0.0002	0.0035	0.0523	0.0058
B*13-C*03	0.0712	0.0391	0.0051	7.6955	0.3857
B*40-C*15	0.0092	0.0039	0.0024	1.6345	0.1143
B*35-DR*12	0.0179	0.0135	0.0070	1.7300	0.6339
B*46-DR*15	0.0150	0.0270	0.0342	0.7900	0.0800

03($\Delta/\delta_{(\Delta)} = 7.69$, $\Delta_{(r)} = 0.38$), 表中用灰色表示. 其余部分表示连锁不平衡程度较低或连锁平衡的单倍型.

3 讨论

如前文所述, HLA 是人类基因组中最具多态性的基因系统, 其重组是产生其遗传多样性的最主要分子基础之一, 也是人类不同种族与不同疾病易感性关联的生物学基础. 深入研究 HLA 重组的分子生物学特性, 对于揭示其等位基因多样性形成的机制以及人体所具有的抵抗各种致病微生物的防御能力有着极其重要的意义.

20 世纪 80 年代起, 国外学者就发现了处于高度连锁不平衡状态的 HLA 复合体发生重组的现象. 研究表明, 整个人类基因组的重组率为 0.92 cM/Mb, 第 6 号染色体的重组率约是 0.71 cM/Mb, 而人类 MHC 的重组率约是 0.49 cM/Mb^[11,12]. 随着分子生物学技术的迅速发展, 微卫星标记技术的逐渐成熟, 很多 HLA 的重组热点被发现, 如 HLA- II 类分子中的 HLA-DNA/RING3 和 DQB3/DQB1, TAP 中 8.8 kb 长度的片段; HLA- I 类中的 HLA-E/P5, HLA-Cw/OTF3, D6S105/HLA-A 和 D6S265/HLA-Cw^[13]. 根据多年来许多科学家在 HLA 重组领域所作的研究, 可以对 HLA 重组特性总结如下: (i) 个体之间的重组率不同; (ii) 重组热点的密度至少是 0.8 Mb 一个; (iii) 个体之间的重组位点的分布也不同; (iv) 除了重组热点之外, 还有重组的温和点 (recombination warm

spots), 即在重组热点之间也会有低水平的重组发生; (v) 重组热点的周围会连锁一些特殊的序列基序, 如一些短片段重复序列, 它们的存在可能会启动 HLA 重组的发生; (vi) 重组率与连锁不平衡的强度成反比; (vii) 女性的重组发生率要比男性高, 但重组热点的模式基本相同^[3,13].

研究表明, HLA 重组并不是随机发生的, 存在着种族特异性和其他一些分子生物学机制, 有待进一步的研究; 而我国目前在此相关领域的研究尚属空白. 本研究在中国汉族人群中发现了 10 例珍贵的 HLA 重组家系, 其中有 8 例都是发生在 HLA- A/Cw 座位之间的重组, 证明了 HLA-A/Cw 确实是重组热点, 与欧裔人种的研究结果一致; 1 例是发生在 HLA-B/DR 之间的重组, 还有 1 例是 HLA-B/Cw 之间的重组, 此类重组在国际上鲜有报道. 发生 HLA-A/Cw 重组的 8 例家系中, 有 4 例存在于 A*02-C*03 单体型, 2 例出现在 A*24-C*03 单体型, 提示这些单体型组合的抗原格局可能更易于 HLA 重组的发生, 但与报道的欧裔重组单体型有所不同. 家系中 B/DR 位点间的重组出现在 DR*02-DQ*05 连锁的单倍型中, 这与国外文献报道的“DR2 很少出现在重组单倍型中”有所差异. 迄今为止, 有关 HLA-B/Cw 位点间的重组国际上仅有一例报道, 这可能与两位点之间的物理图距短 (约 85 kb), 且两位点高度连锁不平衡有着密切的关系. HLA-B 位点是 HLA 复合体中多态性最丰富的基因座位, 这主要是由于重组而造成的; 但是

HLA-B/Cw重组家系的高分辨及基因组测序的结果均表明重组不是发生在HLA-B位点上,而是位于HLA-B和-Cw座位之间^[14]。

我们对随机选择的 1014 例标本的HLA-A、-B、-Cw、-DRB1 及-DQB1 基因座位的低分辨率基因分型资料进行了统计学分析,估算了 16 种重组单体型连锁不平衡度。结果显示,大部分重组单体型连锁不平衡程度都很低,这可能是其容易产生断裂发生重组的主要原因,大多数的学者也都认同重组率与连锁不平衡成反比的观点;但是有 3 个重组单体型却呈强连锁不平衡,即A*33-Cw*03($\Delta/\delta_{(d)} = 7.22$, $\Delta_{(r)} = 0.51$), A*30-Cw*06($\Delta/\delta_{(d)} = 10.37$, $\Delta_{(r)} = 0.93$)和B*13-Cw*03($\Delta/\delta_{(d)} = 7.69$, $\Delta_{(r)} = 0.38$),对其可能的解释有 2 个方面:(i) 这三条单体型对应的另外一条单体型连锁不平衡程度很低,可能在重组中发挥主要作用;(ii) 连锁不平衡与重组率的关系尚不明确,也有文献对两者成绝对反比的关系持反对意见。Termijtelen等人^[15]指出:单体型中发生重组的数量并不比理论上随机发生重组的数量少;同样,由于重组而产生的连锁不平衡单体型也没有明显增加。最近有些研究也发现,在紧密连锁的染色体区域仍存在重组热点,这可能是近期发生的重组事件,因此,在人类重组历史上未留下任何印记。由此可见,人类基因组中可能存在大量的重组热点,而且它们具有一定的流动性^[16]。此外,Kauppi等人^[17]的研究显示,基于连锁不平衡来估计重组热点与单精子作图分析的结果并不完全一致。这就提示连锁不平衡在预测重组热点中的复杂性以及采用实验验证的必要性与重要性。

由此可见,可以解释这种连锁不平衡模式和重组率不一致性的只有自然选择学说。因为连锁不平衡是历史因素造成的,从一定程度上反映了自然选择的作用;而染色体的重组是当即发生的事件,不可能反映一个漫长的选择过程。只有对更多家系样本的 HLA 进行深入研究和利用先进的实验技术,才能最终明确两者之间的关系。本研究发现的 8 个 HLA-A/Cw 重组家系中有 4 例都是发生在 A*02-Cw*03 之间,而通过群体统计学分析,此单体型呈一定程度的连锁不平衡但不是很强($\Delta/\delta_{(d)} = 3.0$, $\Delta_{(r)} = 0.30$)。4 例重组都发生在这样的区域中可以用进化的理论来进行诠释。在随机交配的群体中,连锁不平衡程度随着种群的进化都在不断地下降。比如,或许以前的统计学结论认为 A*02-Cw*03 单体型呈一定的连锁不平衡,

但是随着时间的推移,还是不断会有重组(如本研究中的 4 例重组)发生在此区域,因此有逐渐走向连锁平衡的趋势。因此,单独依赖连锁不平衡程度来衡量重组的发生位点是不够全面和科学的,其分子机制的探讨、两者之间的确切关系和更深入的研究都有待于重组家系的收集、技术方法的进一步更新。

鉴于 HLA 重组的种族特异性,收集更多的中国汉族人群 HLA 重组家系,建立遗传图谱,估算重组频率,探索 HLA 基因重组的发生机制,为研究中华民族的变迁以及 HLA 与疾病的关联提供科学依据和技术平台就显得尤为迫切。

参 考 文 献

- Shiina T, Inoko H, Kulski J K. An update of the HLA genomic region, locus information and disease association. *Tissue Antigens*, 2004, 64: 631—649[DOI]
- 奚永志. MHC 结构与功能. 李幼平, 吴孟超, 郑树森, 主编. 移植免疫生物学. 北京: 高等教育出版社, 2006. 15—38
- Jeffreys A J, Kauppi L, Neumann R. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat Genet*, 2001, 29: 217—222[DOI]
- McVean G, Myers S, Hunt S, et al. The fine-scale structure of recombination rate variation in the human genome. *Science*, 2004, 304: 581—584[DOI]
- Crawford D C, Bhangale T, Li N, et al. Evidence for substantial fine-scale variation in recombination rates across the human genome. *Nat Genet*, 2004, 36: 700—706[DOI]
- Grimwood J, Schmutz J. Genomics: Six is seventh. *Nature*, 2003, 425: 775—776[DOI]
- Kiepiela P, Leslie A J, Honeyborne I, et al. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature*, 2004, 432: 769—775[DOI]
- 孙玉英, 孔繁华, 奚永志, 等. HLA-I类 PCR-SSP 基因分型在异基因造血干细胞移植中的应用. *中华器官移植杂志*, 2001, 22(6): 326—328
- Lee J, Pillai S. Complete nucleotide sequence of MHC class I alleles in the HT29 colon cancer cell line. *Tissue Antigens*, 1993, 42: 530—532
- 谭建明, 周永昌, 唐孝达, 主编. 组织配型技术与临床应用. 北京: 人民卫生出版社, 2002. 62—131
- Cullen M, Perfetto S P, Klitz W, et al. High-resolution patterns of meiotic recombination across the human major histocompatibility complex. *Am J Hum Genet*, 2002, 71: 759—776[DOI]
- Walsh E C, Mather K A, Schaffner S F, et al. An integrated haplotype map of the human major histocompatibility complex. *Am J Hum Genet*, 2003, 73: 580—590[DOI]
- 骆媛, 孙玉英, 奚永志. 人类白细胞抗原分子遗传中的基因重组. *中华医学遗传学杂志*, 2005, 22(4): 427—430
- 骆媛, 孙玉英, 金荔, 等. 一种罕见的人类白细胞抗原 HLA-B/Cw 座位间的基因重组. *中华医学杂志*, 2006, 86(9): 628—631
- Termijtelen A, D'Amaro J, van Rood J J, et al. Linkage disequilibrium in HLA can not be explained by selective recombination. *Tissue Antigens*, 1995, 46: 387—390
- Jeffreys A J, Neumann R, Panayi M, et al. Human recombination hot spots hidden in regions of strong marker association. *Nat Genet*, 2005, 37: 601—606[DOI]
- Kauppi L, Stumpf M P, Jeffreys A J. Localized breakdown in linkage disequilibrium does not always predict sperm crossover hot spots in the human MHC class II region. *Genomics*, 2005, 86: 13—24[DOI]