

研究报告

基于三代测序的食蟹猴 *Mafa-B* 等位基因共表达与进化分析

胡清秀, 黄小琪, 赵爱民, 凌飞*

华南理工大学生物科学与工程学院, 广东 广州 510006

胡清秀, 黄小琪, 赵爱民, 凌飞. 基于三代测序的食蟹猴 *Mafa-B* 等位基因共表达与进化分析[J]. 微生物学通报, 2022, 49(12): 5222-5241

Hu Qingxiu, Huang Xiaoqi, Zhao Aimin, Ling Fei. Co-expression and evolution of *Mafa-B* alleles in *Macaca fascicularis* based on third-generation sequencing[J]. Microbiology China, 2022, 49(12): 5222-5241

摘要:【背景】主要组织相容性复合体(major histocompatibility complex, MHC)的多态性在很大程度上会影响生物医学实验的结果, 而且特定的 *MHC-B* 等位基因与多种疾病的发展进程密切相关。食蟹猴(*Macaca fascicularis*, *Mafa*)是一种开展生物医学研究的重要实验动物, 与人类相比, 目前尚缺乏对食蟹猴 *Mafa-B* 等位基因的综合表征。【目的】获得全面的食蟹猴 *Mafa-B* 等位基因信息, 鉴定 *Mafa-B* 等位基因共表达与进化关系。【方法】基于三代测序获得的食蟹猴 *MHC-B* 基因组信息, 设计特异性引物扩增 33 只越南食蟹猴群体中的 *Mafa-B* 序列, 并结合多种生物信息学方法进行分析。【结果】基于 92 个 *Mafa-B* 等位基因信息, 鉴定了 65 个新的 *Mafa-B* 等位基因。其中, 8 个 *Mafa-B* 等位基因与其他地理来源的食蟹猴群体中已报道的序列相同, 32 个 *Mafa-B* 等位基因与其他猕猴物种中已报道的序列相同。此外, 鉴定了 7 个高频 *Mafa-B* 谱系和 7 对共表达的 *Mafa-B* 等位基因, 并检测到了一个潜在的重组事件。进化分析表明不同地理来源的食蟹猴群体 *Mafa-B* 序列具有很高的相似性。【结论】越南食蟹猴群体中共表达的 *Mafa-B* 等位基因经历了某些抗原的选择, 不同地理来源的食蟹猴群体可能微调其 *Mafa-B* 序列以适应病原体的选择压力, 本文为食蟹猴 MHC 遗传背景研究奠定了基础。

关键词: 食蟹猴; 主要组织相容性复合体; *Mafa-B* 基因; 共表达; 进化

基金项目: 广东省发酵与酶工程重点实验室项目(163194083617178)

Supported by: Guangdong Key Laboratory of Fermentation and Enzyme Engineering Project (163194083617178)

*Corresponding author: E-mail: fling@scut.edu.cn

Received: 2022-07-22; Accepted: 2022-08-15; Published online: 2022-09-08

Co-expression and evolution of *Mafa-B* alleles in *Macaca fascicularis* based on third-generation sequencing

HU Qingxiu, HUANG Xiaoqi, ZHAO Aimin, LING Fei*

School of Biology and Biological Engineering, South China University of Technology, Guangzhou 510006, Guangdong, China

Abstract: [Background] Major histocompatibility complex (MHC) is highly polymorphic, thus influencing the outcome of biomedical experiments. Particularly, specific *MHC-B* alleles are closely related to the development of various diseases. *Macaca fascicularis* (*Mafa*) is an important animal model for conducting biomedical research. Currently, *Mafa-B* alleles in *Mafa* have not been comprehensively characterized. [Objective] To explore the comprehensive information on *Mafa-B* alleles in *Mafa* and to identify co-expression and evolution of *Mafa-B* alleles. [Methods] Through third-generation sequencing, the genomic information of *MHC-B* of *Mafa* was obtained. Then, we designed specific primers to amplify *Mafa-B* alleles in a cohort of 33 Vietnamese *Mafa* individuals and characterized *Mafa-B* alleles with multiple bioinformatics methods. [Results] On the basis of information of 92 *Mafa-B* alleles, we identified 65 novel *Mafa-B* alleles. Among them, 8 alleles were identical to sequences previously reported in *Mafa* of other geographical origins and 32 were also found in macaque populations. In addition, we identified 7 high-frequency *Mafa-B* lineages and 7 pairs of co-expressed *Mafa-B* alleles, and detected one potential recombination event. Evolutionary analysis revealed the high similarity of *Mafa-B* sequences in *Mafa* from different geographical origins. [Conclusion] These co-expressed *Mafa-B* alleles in Vietnamese populations have undergone selection pressure by certain antigens, and *Mafa* populations of different geographical origins may fine-tune the *Mafa-B* sequences in response to pathogens. This study lays a foundation for elucidating the genetic background of MHC in *Mafa*.

Keywords: *Macaca fascicularis*; major histocompatibility complex; *Mafa-B* gene; co-expression; evolution

食蟹猴是一种重要的非人灵长类实验动物, 可用于研究各种人类疾病, 如结核病^[1]、获得性免疫缺陷综合征 (acquired immunodeficiency syndrome, AIDS)^[2]、移植研究^[3]和 COVID-19 疫苗开发等^[4]。

主要组织相容性复合体 (major histocompatibility complex, MHC) 是一个复杂的基因组区域, 其编码的 MHC 分子在免疫防御中扮演着重要的角色。MHC 分为 I 类、II 类和 III 类^[5], I 类和 II 类基因的多态性对疾病易感性有显著影响^[6-7]。而在 MHC I 类分子中, 2 和 3 外显子在整个 MHC I 类基因序列中呈现出最高的

多态性, 其编码的抗原结合位点主要位于 $\alpha 1$ 和 $\alpha 2$ 结构域^[8]。

人类主要组织相容性复合体称为人类白细胞抗原 (human leukocyte antigen, HLA), 恒河猴 (*Macaca mulatta*, *Mamu*) 和食蟹猴 (*Macaca fascicularis*, *Mafa*) 的主要组织相容性复合体分别称为 *Mamu* 和 *Mafa*^[9]。HLA-A、HLA-B 和 HLA-C 基因是单拷贝的, 而 HLA-A 和 HLA-B 基因的同源物在猕猴中是多拷贝的^[10]。截至目前, 在猕猴中尚未发现 HLA-C 的同源物^[11]。人类和猕猴 MHC 基因组的组织结构高度相似, 但是由于多轮复制, 导致猕猴 MHC I 类 A 和 B 基因的数量高度

可变;相较于 *MHC-A* 基因, *MHC-B* 基因具有更复杂的基因数量和可变的转录水平,而且大多数 *MHC-B* 基因是不转录的^[10]。在恒河猴中,一个 *Mamu-B* 单倍型包含 1-6 个主要转录的 *Mamu-B* 基因和 1-10 个次要转录的 *Mamu-B* 基因^[12-13],一个单倍型上同时存在的 *Mamu-B* 基因可多达 19 个^[14]。食蟹猴 *Mafa-B* 单倍型上的基因数量为 1-17 个^[15-19]。一个 *Mafa-B* 单倍型可由 1-7 个主要转录基因和 1-15 个次要转录基因组成^[15-19]。猕猴中 *MHC-B* 基因由于广泛的基因复制而变得复杂,目前将不同的 *MHC-B* 等位基因分配到单倍型上特定的 *B* 基因座尤为困难。

人类 *MHC* 基因已被深入研究,8 849 个 *HLA-B* 等位基因(IPD-HLA Release 3.48.0)被报道,而食蟹猴中仅有 1 090 个 *Mafa-B* 等位基因(IPD-MHC Release 3.9.0.0)被报道。与人类相比, *Mafa-B* 等位基因的全面信息仍然缺乏。据报道,某些特定的 *MHC-B* 等位基因与多种疾病的进展有关。例如, *Mamu-B*001* 阳性个体对胶原诱导性关节炎(collagen-induced arthritis, CIA)具有抗性^[20]; *HLA-B*35* 和 *HLA-B*58* 与人类免疫缺陷病毒(human immunodeficiency virus, HIV)的易感性有关^[21];而携带 *HLA-B*57* 和 *HLA-B*27* 的个体在感染 HIV 后呈现出较低的病毒载量和缓慢的疾病进展^[22]。此外,在感染了猴免疫缺陷病毒(simian immunodeficiency virus, SIV)的恒河猴中,携带 *Mamu-B*08* 或 *Mamu-B*017* 的个体病毒载量相对较低,疾病进展缓慢^[23-24]。一些 *MHC-B* 分子的肽结合基序在不同物种中是相同的。例如, *Mamu-B*039:01* 和越南来源的 *Mafa-B*039:01* 中都含有一个特定的肽结合基序^[25]; *Mamu-B*03* 和 *Mamu-B*08* 的肽结合基序与 *HLA-B*27* 的肽结合基序非常相似^[26-27],它们可能执行相同的功能。研究表明, *MHC* 分子

和杀伤细胞免疫球蛋白样受体(killer cell immunoglobulin-like receptor, KIR)之间的相互作用与 AIDS 进展有关。例如, *KIR3DL05/KIR3DS05/KIR3DL10* 和 *Mamu-B*012* 基因组合显示出高病毒载量,疾病进展迅速^[28]。这表明,准确有效地表征非人灵长类动物模型 *MHC-B* 基因的遗传进化特征可能对未来疫苗设计具有重要意义。

本研究基于三代测序获得的食蟹猴 *MHC-B* 基因组序列,利用基因特异性扩增技术,通过对 *Mafa-B* 基因的 2 和 3 外显子进行测序,获得了 33 只越南食蟹猴群体中的 *Mafa-B* 等位基因序列。基于群体中的 *Mafa-B* 等位基因序列,进一步对 *Mafa-B* 基因进行共表达和进化分析。本研究提供了越南来源食蟹猴 *Mafa-B* 等位基因特定组合和遗传进化的关键信息,以期为未来的生物医学和进化研究提供重要参考,同时也为其他重要免疫基因组的特征研究提供资源,比如 *KIR* 基因组。

1 材料与方法

1.1 材料

1.1.1 实验动物

研究使用的 33 只健康越南食蟹猴来自于广东省生物资源应用研究所华南灵长类研究开发中心。每只猴子取 3-5 mL 的外周血样本,保存在经 EDTA 处理的真空管中。实验经广东蓝岛生物技术有限公司实验动物管理与使用委员会(Institutional Animal Care and Use Committee, IACUC)审批通过(项目编号: IACUC-003)。

1.1.2 主要试剂、仪器和培养基

E.Z.N.A.TM Blood RNA Kit, Omega 公司; PrimeScriptTM II 1st Strand cDNA Synthesis Kit、DL2000 DNA Marker 和 pMDTM19-T 载体,

TaKaRa 公司; Green Taq Mix, Vazyme 公司; 大肠杆菌 DH5 α 感受态细胞, 生工生物工程(上海)股份有限公司; DNA 凝胶微量回收纯化试剂盒, Magen 公司; 胰蛋白胨、酵母提取物、NaCl 和琼脂粉, OXOID 公司; 氨苄青霉素钠, 阿拉丁公司。DNA 测序仪, 广州擎科公司; PCR 扩增仪和电泳凝胶成像仪, Bio-Rad 公司。

LB 液体培养基(g/L): 胰蛋白胨 10.0, 酵母提取物 5.0, NaCl 10.0。LB 固体培养基(g/L): 在 LB 液体培养基中加入 15.0 g/L 的琼脂粉。培养基均于 1×10^5 Pa 灭菌 30 min。LB 氨苄抗性培养基: 添加初始浓度为 100 mg/mL 的氨苄青霉素钠^[29]。

1.2 方法

1.2.1 食蟹猴 *MHC-B* 基因区域的识别和注释

首先, 使用 BLAST(V2.2.26) 将从免疫多态性数据库(Immuno Polymorphism Database, IPD) 下载的 *MHC-B* 基因(食蟹猴和恒河猴)的编码序列(coding sequence, CDS), 与本课题组三代测序组装的食蟹猴基因组(GenBank 登录号为 PRJNA819149) 进行比对, 从原始比对结果中保留最优比对结果, 从而确定 *MHC-B* 基因区域的位置。其次, 将比对得到的食蟹猴 *MHC* 基因组序列与 IPD 中公布的 *MHC-B* 等位基因序列进行比对, 精细注释出每个 *MHC-B* 基因的外显子和内含子。最后, 将注释得到的 *MHC-B* 基因序列提交至 GenBank 获得序列号, 并提交至国际非人灵长类动物 MHC 命名委员会命名。

1.2.2 RNA 提取、cDNA 克隆和测序

使用 E.Z.N.A.TM Blood RNA Kit 从 33 只健康的越南食蟹猴外周血白细胞中提取总 RNA, 接着使用 PrimeScriptTM II 1st Strand cDNA Synthesis Kit 合成 cDNA。我们根据 10 个 *Mafa-B* 基因的碱基序列差异性设计了 10 对特异性引物(表 1) 扩增 *Mafa-B* 基因的 2 和 3 外显子序列。

PCR 反应体系(50 μ L): Green Taq Mix 25 μ L, 上、下游引物(10 μ mol/L) 各 2 μ L, cDNA 模板(100 ng/ μ L) 1 μ L, ddH₂O 20 μ L。PCR 反应条件: 95 $^{\circ}$ C 5 min; 95 $^{\circ}$ C 30 s, 58–64 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 25–50 s, 34 个循环; 72 $^{\circ}$ C 10 min; 4 $^{\circ}$ C 保存。由于 Taq DNA 聚合酶无 3'→5' 核酸内切酶活性, 引物 3' 端最后一个碱基与模板是否匹配决定着能否扩增出目的产物, 因此可以针对不同基因中个别碱基差异而进行选择性的扩增^[30]。PCR 产物用 DNA 凝胶微量回收试剂盒纯化后与 pMDTM19-T 载体连接, 连接后的产物转化大肠杆菌 DH5 α 感受态细胞中, 在氨苄抗性(终浓度为 100 μ g/mL) 的 LB 固体培养基上 37 $^{\circ}$ C 培养 12–16 h 后, 挑选单克隆至氨苄抗性(终浓度 100 μ g/mL) 的 LB 液体培养基中, 30 $^{\circ}$ C、200 r/min 培养过夜。使用菌液 PCR 的方法挑选阳性克隆子, 菌液 PCR 的引物为通用引物 M13-F (5'-GTA AACGACGGCCAGT-3') 和 M13-R (5'-CAGG AACAGCTATGAC-3')。每个转化平板中选择 10–50 个阳性克隆子送至广州擎科公司进行测序。使用 SnapGene 4.1.9 和 MEGA 11 软件分析和比对序列, 当 1 个序列至少在 3 个独立克隆中相同时可以被确定为 1 个等位基因序列。所有等位基因序列提交到 GenBank 获取序列号, 并与 IPD 公开的等位基因进行 BLAST 比对确定新等位基因, 新等位基因提交至国际非人灵长类动物 MHC 命名委员会命名。

1.2.3 *Mafa-B* 等位基因进化分析和重组分析

使用 MEGA 11 中的 Clustal W 程序比对序列, 采用邻接法和 p-distance 模型构建进化树, bootstrap 参数设置为 1 000。使用重组检测程序(recombination detection program version 4, RDP4) 进行重组分析, 推断重组事件的最高可接受 *P* 值设置为 0.000 005, 窗口大小设置为 20 个核苷酸, 使用至少 4 种检测方法确定的重组事

表 1 特异性扩增 *Mafa-B* 基因的引物Table 1 Primers for specific amplification of *Mafa-B* genes

基因 Gene	引物名称 Primer name	序列 Sequence (5'→3')	退火温度 Annealing temperature (°C)	延伸时间 Extension time (s)	产物长度 Product length (bp)
<i>B1</i>	F1	ACGCAGTTTGTGCGTTTCGACAGCA	58	47	847
	R1	AGCTCTTCCTCCTCCACATCACAGCG			
<i>B2</i>	F2	ATGCGGGTTATGGCGCCCGGA	64	25	498
	R2	GTTCTGAGCCGCCTCACCCGCA			
<i>B3</i>	F3	TCGCACTCCATGAGGTATTTTCAGCAA	62	50	850
	R3	CGATGGGGATGGTGGACTGGGT			
<i>B4</i>	F4	GCTCGCACTCCTTGAGGTATTTCT	58	34	616
	R4	GCACCTCAGGGTGGCCTCATGGTT			
<i>B5</i>	F5	CCGTGTCGGCCCGCCGCGGC	60	35	663
	R5	CACAAGTTCGGTGCCTGAGTTTGA			
<i>B6</i>	F6	CCCACTCCATGAGGTATTTCC	60	35	673
	R6	TGAGTTTGCTCCTCCCCATCT			
<i>B7</i>	F7	ATGATGCCCCGAACCGTCCT	60	40	665
	R7	CAGGGTGGCCTCATGGTCAGAT			
<i>B8</i>	F8	CCTGGCCCTGACCGAGACCTGGGT	58	35	616
	R8	ATGGTCAGAGACGGGGTGGTGGGC			
<i>B9</i>	F9	GCGAACCCTCCTCCTGCTGCTCTCA	62	30	592
	R9	CAGCGTCTCCTCCCGTTCTCCAT			
<i>B10</i>	F10	GCACCTCAGTGTCGGCCT	58	27	508
	R10	TGCAGCGTCTCCTTCCTGTTCTCT			

件被认为是有效的；使用重组识别程序 (recombination identification program, RIP) 进行重组事件的验证 (<http://www.hiv.lanl.gov/>)，窗口大小设置为 200，置信区间设置为 99.0%^[25]。

2 结果与分析

2.1 基于三代测序的食蟹猴 *MHC-B* 基因注释良好

基于课题组三代测序组装得到的食蟹猴基因组，contig utg0003481 (总长度 8 094 345 bp) 上 4.9–5.4 Mb 范围与食蟹猴和恒河猴的 *MHC-B* 等位基因的 CDS 序列共线性良好 (图 1)，序列一

致性为 93.7%–100.0%。食蟹猴 *MHC-B* 基因区域长度约为 500 kb，包含 10 个 *B* 基因 (GenBank 登录号为 MZ254652–MZ254661)。按照在染色体上的顺序将这 10 个 *B* 基因分别编号为 *B1–B10*，并提交至国际非人灵长类动物 *MHC* 命名委员会获得官方命名 (表 2)。

2.2 *Mafa-B* 等位基因的鉴定和频率

在 33 只越南来源食蟹猴群体中进行了 *Mafa-B* 等位基因特异性扩增，结果发现，用 10 个 *Mafa-B* 基因特异性引物进行扩增后产物大小与理论值符合，表明扩增成功 (图 2)。由于使用通用引物 M13-F/M13-R 进行了阳性克隆鉴

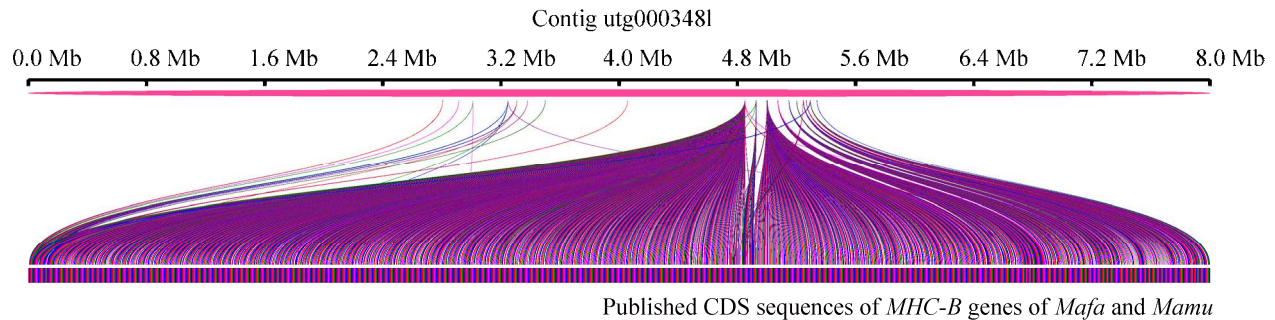


图 1 Contig utg0003481 与食蟹猴和恒河猴 *MHC-B* 基因的 CDS 序列共线性 图的顶部为 contig utg0003481 的序列, 用粉红色表示; 刻度线上的数字代表 contig utg0003481 的序列长度; 图的底部为从 IPD 下载的食蟹猴和恒河猴已发表的 *MHC-B* 基因的 CDS 序列, 用紫色表示

Figure 1 Contig utg0003481 showed good synteny with CDS sequences of *MHC-B* genes of *Mafa* and *Mamu*. The sequence of contig utg0003481 is shown in pink at the top of the graph; The numbers on the scale represent the sequence length of contig utg0003481; The published CDS sequences of *MHC-B* genes of *Mafa* and *Mamu* downloaded from IPD are shown in purple at the bottom of the graph.

表 2 10 个 *Mafa-B* 基因的官方命名

Table 2 Official designations of 10 *Mafa-B* genes

序列号	基因	官方命名
Accession No.	Gene	Official designation
MZ254661	<i>Mafa-B1</i>	<i>Mafa-B*162:01:02:01</i>
MZ254660	<i>Mafa-B2</i>	<i>Mafa-I*01:57:01:01</i>
MZ254659	<i>Mafa-B3</i>	<i>Mafa-B*056:05:01:01</i>
MZ254658	<i>Mafa-B4</i>	<i>Mafa-B*109:30:01:01</i>
MZ254657	<i>Mafa-B5</i>	<i>Mafa-B*034:04:01:01</i>
MZ254656	<i>Mafa-B6</i>	<i>Mafa-B11L*01:07:01:01</i>
MZ254655	<i>Mafa-B7</i>	<i>Mafa-B16*01:03:02:01</i>
MZ254654	<i>Mafa-B8</i>	<i>Mafa-B17*01:06:01:01</i>
MZ254653	<i>Mafa-B9</i>	<i>Mafa-B*054:04:01:01</i>
MZ254652	<i>Mafa-B10</i>	<i>Mafa-B19Ps*01:01:01:02</i>

定, 阳性克隆产物长度要加上 pMDTM19-T 载体上大约 111 bp 的长度, 结果表明阳性克隆产物大小与理论值符合(图 3)。

研究共测序了 6 437 个阳性克隆子, 经测序后进行序列分析比对, 获得了 5 351 条 cDNA 序列。接着过滤掉少于 3 个阳性克隆中检测到的相同序列, 确定了 92 个 *Mafa-B* 等位基因序列 (GenBank 登录号 OK486180–OK486272)。其中,

65 个等位基因代表新的 *Mafa-B* 等位基因, 其余 27 个等位基因在食蟹猴群体中已有报道。特异性扩增结果表明, *B2*、*B3*、*B5* 和 *B6* 基因特异性引物扩增出了代表性的 *Mafa-B* 谱系等位基因 (*I*01*、*B*056*、*B*034*、*B11L*); *B7* 和 *B9* 基因特异性引物只扩增出了少量等位基因 (*B*079*、*B*017*、*B*180*、*B*124* 和 *B*021*) (表 3); 而用 *B1*、*B4*、*B8* 和 *B10* 基因特异性引物扩增后未鉴定到等位基因(未鉴定到 3 个及 3 个以上相同的阳性克隆序列, 所以不能确定为等位基因)。研究表明, *B1*、*B4* 和 *B9* 基因的转录水平很低^[15-19], 而 *B7*、*B8* 和 *B10* 基因被鉴定是假基因^[10]。*MHC* 基因以共显性方式表达, 杂合子个体在一个位点可以扩增出 2 种 *Mafa-B* 等位基因。

在分析的 33 只越南来源食蟹猴中, 每个个体表达 4–8 个 *Mafa-B* 基因。其中, 15 个个体表达了 4 个 *Mafa-B* 基因, 13 个个体表达了 5 个 *Mafa-B* 基因, 4 个个体表达了 6 个 *Mafa-B* 基因, 1 个个体表达了 8 个 *Mafa-B* 基因(表 3)。这些数据表明, *Mafa-B* 基因在食蟹猴群体中是高度重复的。经统计, 分布频率大于 10% 的 *Mafa-B* 等

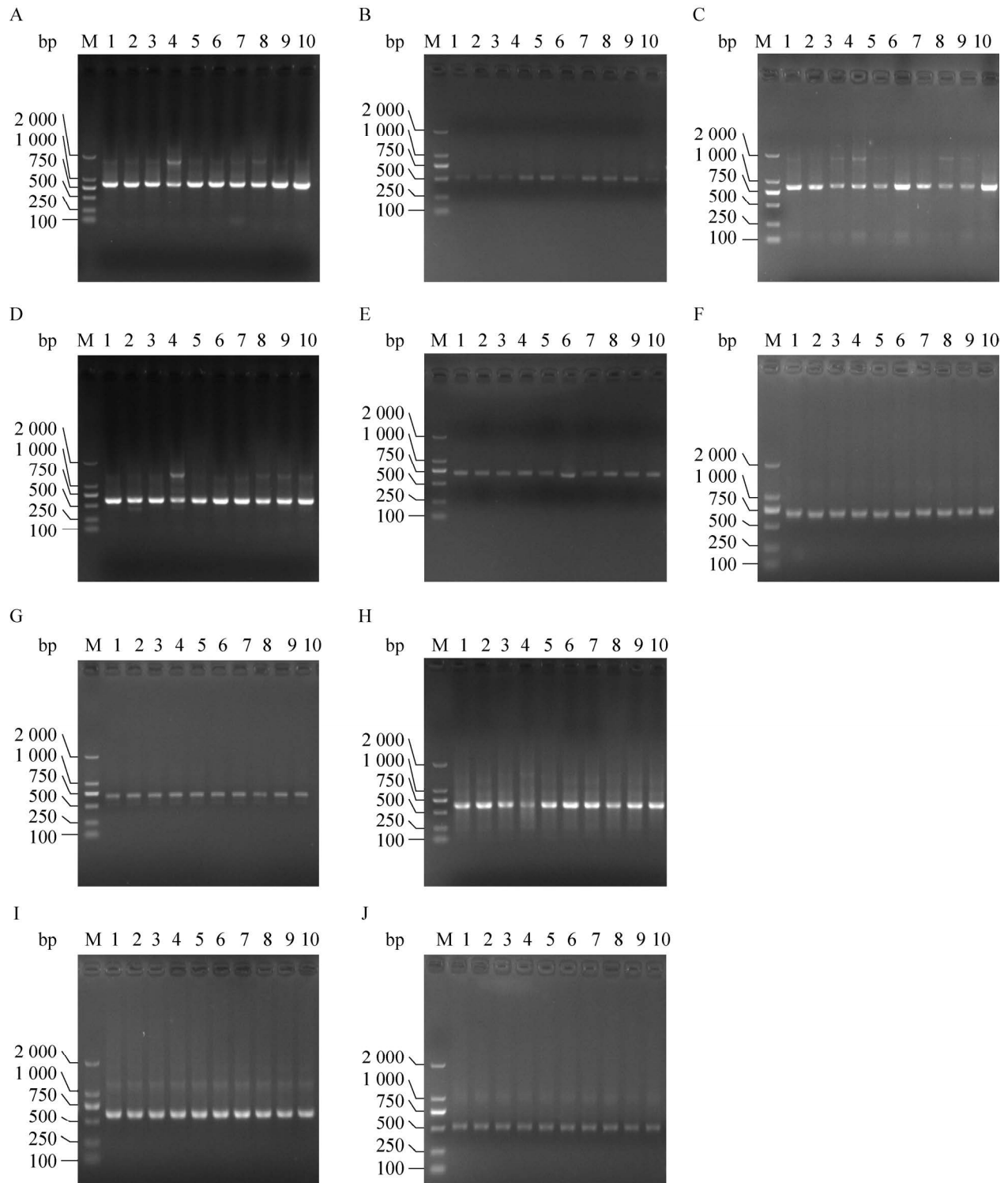


图 2 10 个 *Mafa-B* 基因的 PCR 扩增结果 A-J: *B1-B10*; M: DL2000 DNA Marker; 1-10: 1-10 号食蟹猴个体的特异性扩增产物

Figure 2 Results of PCR amplification of 10 *Mafa-B* genes. A-J: *B1-B10*; M: DL2000 DNA Marker; 1-10: Specific amplification products of *Mafa* No. 1-10.

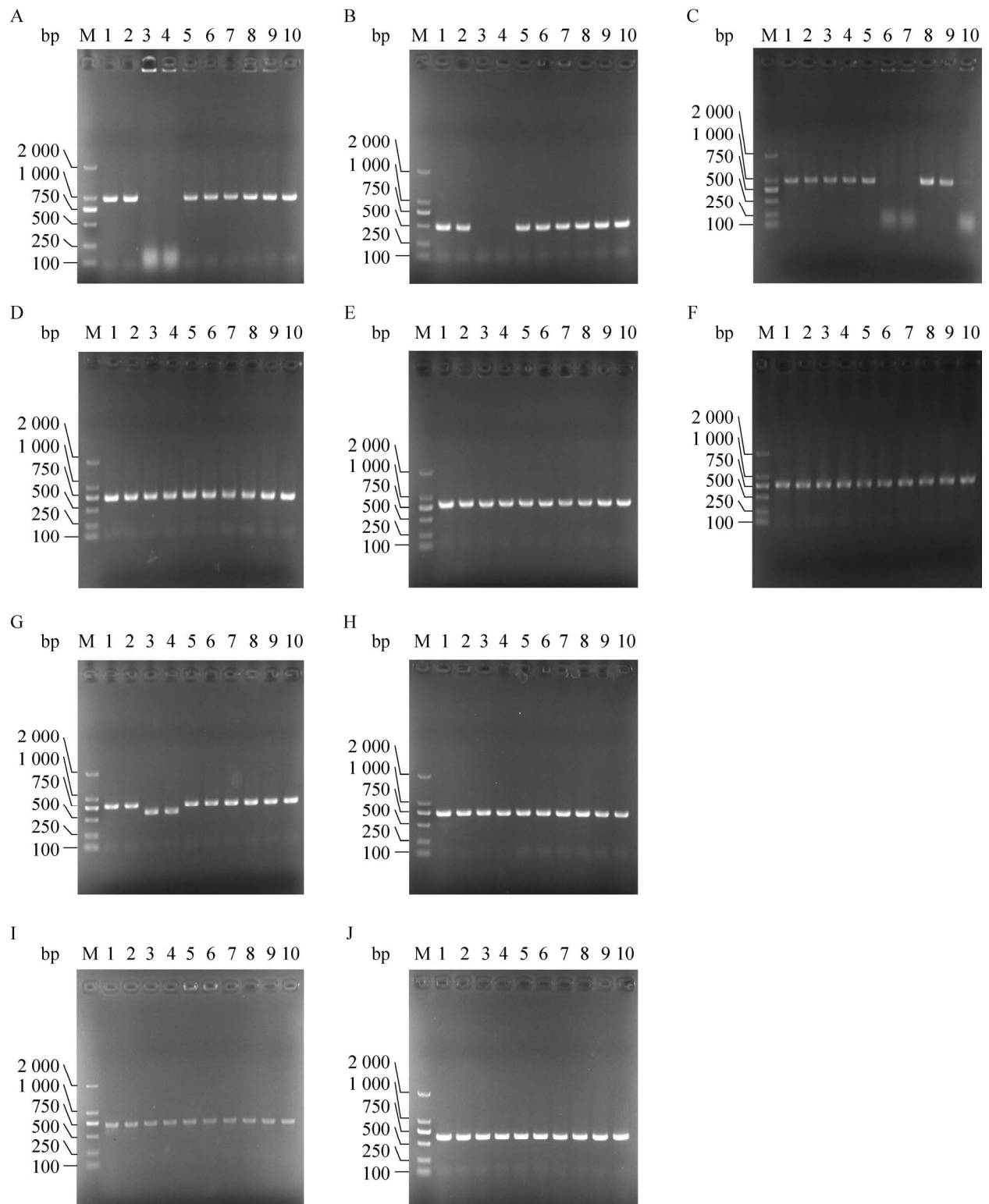


图 3 阳性克隆鉴定 A-J: *BI-B10*; M: DL2000 DNA Marker; 1-10: 1-10 号菌的扩增产物

Figure 3 Identification of positive clones. A-J: *BI-B10*; M: DL2000 DNA Marker; 1-10: Amplification products of bacteria No. 1-10.

表 3 *Mafa-B* 等位基因在 33 只食蟹猴群体中的分布
Table 3 Distribution of *Mafa-B* alleles in 33 *Mafa* individuals

动物编号 Animal No.	等位基因 Alleles	B3	B5	B6	B7	B9
1	<i>I*01:14:05:02</i>	<i>B*056:05:01:01</i>	<i>B*034:04:01:01</i>	<i>B*044:01:01:03nov</i>		<i>B*180:02:01:02nov</i>
2	<i>B*072:03nov1</i>	<i>B*056:05:01:01</i>	<i>B*034:04:01:01</i>	<i>B*089:01:05nov</i> <i>B*021:02nov</i>		<i>B*124:03:01:01nov</i>
3	<i>I*01:31:01:01</i>	<i>B*056:01:01:01</i>	<i>B*007:10nov</i>	<i>B*082:08:01:01nov</i>		
4	<i>I*01:43:03:01</i>	<i>B*028:05nov</i>	<i>B*004:04:01:01</i>	<i>B*011:07nov</i>	<i>B*079:06:01:01</i>	<i>B*021:07nov</i>
5	<i>B*072:12:02:01nov</i>	<i>B*068:02:01:02nov</i>				
6	<i>I*01:14:05:02</i>	<i>B*056:05:01:01</i>	<i>B*034:04:01:01</i>	<i>B*045:03:02:01nov</i>		
7	<i>I*01:43:03:01</i>	<i>B*056:05:02nov2</i>	<i>B*034:04:01:01</i>	<i>B*137:06:01:01nov</i>		
8	<i>I*01:43:03:01</i>	<i>B*007:09</i>	<i>B*007:01:06:01nov</i>	<i>B*001:01:01:02</i>		
9	<i>B*149:01:01:02nov</i>	<i>B*056:01:01:01</i>	<i>B*034:04:01:01</i>	<i>B*101:08:01:01</i> <i>B11L*01:01nov1</i>	<i>B*079:10:01:01N</i>	
10	<i>I*01:43:03:01nov</i>	<i>B*007:09</i>	<i>B*007:01:06:01nov</i>	<i>B*001:01:01:02</i> <i>B11L*01:06:01:01Nnov1</i>		
11	<i>B*149:02:01:01nov</i>	<i>B*056:01:01:01</i>	<i>B*105:01nov1</i>	<i>B*156:01:01:01nov</i>		
12	<i>B*098:04:01:02nov</i>	<i>B*041:02:01:03</i>	<i>B*013:10:01:01</i>	<i>B*001:01:01:02</i>	<i>B*079:10:01:01N</i>	
13	<i>I*01:43:03:01nov</i>	<i>B*056:06</i>	<i>B*066:01:01:01</i>	<i>B*038:01:02:01nov</i> <i>B11L*01:01nov2</i>		
14	<i>B*098:16:02:01nov</i>	<i>B*007:09</i>	<i>B*019:05</i>	<i>B*082:06nov</i>		
15	<i>B*036:05</i>	<i>B*068:02:01:02nov</i>	<i>B*007:01:06:01nov</i>	<i>B*045:07:01:01</i> <i>B*082:05nov</i>		
16	<i>I*01:70</i>	<i>B*068:06:01:03nov</i>	<i>B*032:02:01:01nov</i>	<i>B*001:01:01:02</i>	<i>B*079:09:01:01</i>	
17	<i>I*01:43:03:01nov</i>	<i>B*028:04:01:01nov</i> <i>B*068:02:01:02nov</i>	<i>B*034:04:01:01</i>	<i>B11L*01:06:01:01Nnov2</i>		
18	<i>I*01:43:03:01nov</i>	<i>B*144:01:01:01nov</i>	<i>B*034:04:01:01</i>	<i>B*082:08:01:01nov</i>		
19	<i>I*01:43:03:01nov</i>	<i>B*056:05:02nov1</i>	<i>B*034:04:01:01</i>	<i>B*116:01nov</i>	<i>B*017:01</i>	

(待续)

(续表 3)

19	<i>B*098:04:01:02nov</i>	<i>B*144:02:01:01nov</i>	<i>B*104:01:02:01nov</i>	<i>B*057:02:01:01nov</i>
20	<i>I*01:43:03:01nov</i>	<i>B*007:09</i>	<i>B*007:10nov</i>	<i>B*039:01:02nov</i> <i>B*079:02:03:02</i>
21	<i>I*01:43:03:01nov</i>	<i>B*056:05:02nov2</i>	<i>B*034:04:01:01</i>	<i>B*045:01:01:01nov</i>
22	<i>I*01:43:03:01nov</i>	<i>B*056:05:02nov2</i>	<i>B*034:03:01:01nov</i>	<i>B*044:01:01:03nov</i>
23	<i>I*01:43:03:01nov</i>	<i>B*056:05:02nov2</i>	<i>B*081:05:02:01nov</i>	<i>B*161:05:01:01nov</i> <i>B*079:10:01:01N</i>
24	<i>B*072:19:01:01nov</i>	<i>B*007:09</i>	<i>B*018:01:01:02nov</i>	<i>B11L*01:06:01:01Nnov3</i> <i>B*017:02:01:01</i>
25	<i>I*01:35:03:01</i>	<i>B*068:06:01:03nov</i>	<i>B*007:01:06:01nov</i>	<i>B11L*01:05nov</i>
26	<i>I*01:43:03:01</i>	<i>B*041:02:01:03</i>	<i>B*145:01:01:01</i>	<i>B*137:03:01:01nov</i>
27	<i>I*01:14:08:02</i>	<i>B*069:04nov</i>	<i>B*002:04:01:03nov1</i>	<i>B*082:06nov</i>
28	<i>I*01:43:03:01nov</i>	<i>B*028:04:01:01nov</i>	<i>B*013:11:01:01</i>	<i>B11L*01:06:01:01Nnov2</i>
29	<i>B*072:08nov</i>	<i>B*048:05nov</i>	<i>B*002:04:01:03nov2</i>	<i>B*001:01:01:02</i>
30	<i>B*072:18nov</i>	<i>B*056:05:02nov2</i>	<i>B*002:04:01:03nov1</i>	<i>B*001:01:01:02nov1</i>
31	<i>I*01:43:03:01</i>	<i>B*101:08:01:01</i>	<i>B*004:04:01:01</i>	<i>B*011:07nov</i>
32	<i>B*098:16:02:01nov</i>	<i>B*075:04:01:01</i>	<i>B*108:02:01:01nov</i>	<i>B*001:01:01:02nov1</i>
33	<i>B*072:03nov1</i>	<i>B*056:05:02nov2</i>	<i>B*105:01:01:01nov2</i>	<i>B*001:01:01:02nov2</i>

注: 官方命名的等位基因用粗体标记; 名称带有“nov”的等位基因只包含部分 2 和 3 外显子序列, 不符合 IPD-MHC 数据库的最小序列长度标准, 因此不能被官方命名; 某些序列与 IPD-MHC 数据库中同一等位基因的一致性最高, 在这种情况下, “nov”后面会加数字(比如“nov1”, “nov2”和“nov3”)来区分这些不同的序列

Note: Alleles highlighted in bold are official names; Alleles with “nov” in the name only contain partial exons 2 and 3 sequences which could not be officially named as they do not meet the minimal sequence length criteria of the IPD-MHC database; Some sequences were found to show the highest identity to the same allele in the IPD-MHC database, in these cases, “nov” is followed by a number (such as “nov1”, “nov2” and “nov3”) to distinguish between these different sequences.

位基因谱系有 *I*01:43* (16/33, 48.5%)、*B*056:05* (10/33, 30.3%)、*B*007:09* (5/33, 15.2%)、*B*034:04* (10/33, 30.3%)、*B*007:01* (5/33, 15.2%)、*B*001:01* (8/33, 24.2%) 和 *B11L*01* (7/33, 21.2%)。这一现象表明, 这 7 个谱系的等位基因可能以高频率分布在群体中。同时, 本研究鉴定到 3 个可能与疾病进展相关的同源基因 (*Mafa-B*056:01*、*Mafa-B*017:02* 和 *Mafa-B*001*)。Mamu-*AI*001:01* 是抗 SIV 感染的保护性等位基因^[31], Mamu-*AI*001:01* 和 Mamu-*B*056:01* 呈递多肽 Gag QI9, 而且具有相同的肽结合关键的 F 口袋残基^[32], *Mafa-B*056:01* 与 Mamu *B*056:01* 在序列上高度相似。Mamu-*B*017:01* 是抗 SIV 感染的保护性等位基因^[24], 而 *Mafa-B*017:02* 与 Mamu-*B*017:01* 的序列高度相似。Mamu-*B*001* 被证明是胶原诱导性关节炎的保护性等位基因^[20], *Mafa-B*001* 与 Mamu-*B*001* 是同源基因。因此, *Mafa-B*056:01*、*Mafa-B*017:02* 和 *Mafa-B*001* 可能与它们的同源基因具有类似的功能。

2.3 共表达 *Mafa-B* 等位基因的鉴定

共表达的 *MHC-B* 等位基因可能与疾病的发生率密切相关。结果表明, 等位基因组合 *B*056:05:02nov2-B*034:04:01:01* (6 和 21 号个体)、*I*01:43:03:01-B*034:04:01:01* (6 和 7 号个体) 和 *B*001:01:01:02-B*079:10:01:01N* (9 和 11 号个体) 频率为 6.1%; 等位基因组合 *B*056:05:01:01-B*034:04:01:01* (1、2 和 5 号个体) 和 *I*01:43:03:01nov-B*056:05:02nov2* (21、22 和 23 号个体) 频率为 9.1%; 等位基因组合 *B*007:09-B*007:01:06:01nov* (7、9、13 和 24 号个体) 和 *I*01:43:03:01nov-B*034:04:01:01* (16、17、18 和 21 号个体) 频率为 12.1% (表 3)。这些共表达的 *Mafa-B* 等位基因表明, 越南食蟹猴群体经历了某些抗原的选择。

2.4 与其他地理来源食蟹猴群体或其他猕猴物种序列相同的 *Mafa-B* 等位基因

在 27 个已报道的等位基因中, 2 个等位基因 (*Mafa-B*145:01:01:01* 和 *Mafa-B*056:01:01:01*) 的 2 和 3 外显子编码的氨基酸序列与毛里求斯食蟹猴序列一致^[33]。6 个等位基因 (*Mafa-B*079:08*、*Mafa-B*079:09:01:01*、*Mafa-B*079:02:03:02*、*Mafa-B*056:05:01:01*、*Mafa-B*041:02:01:03* 和 *Mafa-B*017:01*) 的 2 和 3 外显子编码的氨基酸序列与菲律宾食蟹猴序列一致^[34-35]。32 个 *Mafa-B* 等位基因的 2 和 3 外显子序列与其他猕猴物种的序列相同。其中, 28 个 *Mafa-B* 序列与恒河猴序列一致性为 100% (表 4)。8、15 和 8 个 *Mafa-B* 序列分别与北方豚尾猴 (*Macaca leonine*, Malo)、南方豚尾猴 (*Macaca nemestrina*, Mane) 和熊猴 (*Macaca assamensis*, Maas) 序列一致性为 100% (表 4)。17 个等位基因序列与 ≥3 个猕猴种群中的序列相同。例如, 5 个等位基因 (*Mafa-B*056:05:01:01*、*Mafa-B*056:05:02nov2*、*Mafa-I*01:43:03:01*、*Mafa-I*01:43:03:01nov* 和 *Mafa-I*01:14:08:02*) 与食蟹猴、恒河猴、南方豚尾猴和北方豚尾猴群体中的序列相同 (表 4)。大多数序列是越南食蟹猴特有的。

2.5 *Mafa-B* 等位基因进化分析

基于序列的相似性, 对本研究鉴定的 92 个越南来源 *Mafa-B* 序列与印度尼西亚^[18]、菲律宾^[34-35] 和毛里求斯^[33] 来源的 61 个 *Mafa-B* 序列之间的进化关系进行分析。一些不同地理来源的 *Mafa-B* 序列处于同一个分支上 (图 4), 表明这些 *Mafa-B* 序列之间具有很高的相似性。同时, 我们鉴定了 12 组基因对, 其中一个序列在越南种群中发现, 另外一个序列在越南以外的种群发现 (图 4 中粗体)。其中, 7、3 和 2 组中的另一个序列分别来自菲律宾、印度尼西亚和毛里求斯。虽然大部分 *Mafa-B* 序列是特定地域所特

表 4 与其他猕猴物种相同的 *Mafa-B* 等位基因的总结Table 4 Summary of *Mafa-B* alleles identical to other macaque species

序列号 Accession No.	等位基因 Allele	与其他猕猴物种相同的等位基因 Alleles identical to other macaque species
OK486184	<i>Mafa-B*079:08</i>	<i>Mamu-B*079:01:01:01</i> (EF611165) <i>Mane-B*079:01:01</i> (GQ153468)
OK486185	<i>Mafa-B*079:09:01:01</i>	<i>Mamu-B*079:01:02:02</i> (LR990716) <i>Mane-B*079:01:07</i> (HQ110970)
OK486187	<i>Mafa-B*079:02:03:02</i>	<i>Mamu-B*079:01:02:02</i> (LR990716) <i>Mane-B*079:01:07</i> (HQ110970)
OK486188	<i>Mafa-B*017:02:01:01</i>	<i>Mamu-B*017:02:02</i> (LR655877) <i>Malo-B*017:01</i> (KT214461)
OK486189	<i>Mafa-B*044:01:01:03nov</i>	<i>Mamu-B*44*08</i> (NHP06746)
OK486196	<i>Mafa-B*001:01:01:02</i>	<i>Mamu-B*001:01:01:02</i> (LR743803)
OK486202	<i>Mafa-B11L*01:01nov2</i>	<i>Mamu-B11L*01:05:01:01</i> (EU429632)
OK486218	<i>Mafa-B*007:10nov</i>	<i>Mamu-B*007:04:02</i> (AB540183) <i>Mane-B*007:02</i> (KF012964)
OK486219	<i>Mafa-B*004:04:01:01</i>	<i>Mamu-B*004:04:01:01</i> (NHP02221)
OK486220	<i>Mafa-B*007:01:06:01nov</i>	<i>Mamu-B*007:02:02</i> (LM608024) <i>Maas-B*007:01</i> (NHP05201)
OK486221	<i>Mafa-B*105:01:01:01nov1</i>	<i>Mamu-B*105:01</i> (AB540184) <i>Mane-B*105:02</i> (GQ131759)
OK486223	<i>Mafa-B*066:01:01:01</i>	<i>Mamu-B*066:01:01:01</i> (EU682526)
OK486226	<i>Mafa-B*104:01:02:01nov</i>	<i>Mane-B*104:02</i> (GQ131759)
OK486228	<i>Mafa-B*081:05:02:01nov</i>	<i>Mamu-B*081:01:01:01</i> (EF611167) <i>Mane-B*081:02</i> (LN899612)
OK486230	<i>Mafa-B*018:01:01:02nov</i>	<i>Mamu-B*018:01</i> (AM902534) <i>Malo-B*018:01</i> (KT214460)
OK486232	<i>Mafa-B*002:04:01:03nov1</i>	<i>Mamu-B*002:03</i> (LT546138)
OK486233	<i>Mafa-B*013:11:01:01</i>	<i>Malo-B*013:01</i> (KT214463)
OK486234	<i>Mafa-B*002:04:01:03nov2</i>	<i>Mamu-B*002:03</i> (LT546138)
OK486237	<i>Mafa-B*056:05:01:01</i>	<i>Mamu-B*056:02:01:04</i> (LR743859) <i>Mane-B*056:01</i> (FJ875237) <i>Malo-B*056:01</i> (KT214446)
OK486240	<i>Mafa-B*056:05:02nov2</i>	<i>Mamu-B*056:02:01:04</i> (LR743859) <i>Malo-B*056:01</i> (KT214446) <i>Mane-B*056:01</i> (FJ875237)
OK486242	<i>Mafa-B*056:01:01:01</i>	<i>Mamu-B*056:01:01:02</i> (LR743854)
OK486243	<i>Mafa-B*041:02:01:03</i>	<i>Mamu-B*041:01:01:02</i> (LR723148)
OK486244	<i>Mafa-B*056:06</i>	<i>Mamu-B*056:03:01:01</i> (AB540195)
OK486252	<i>Mafa-B*101:08:01:01</i>	<i>Mamu-B*101:06:01</i> (LR723080)
OK486254	<i>Mafa-I*01:14:05:02</i>	<i>Mamu-I*01:52:01:01</i> (LT908865) <i>Mane-I*01:06:03</i> (LN875482)

(待续)

(续表 4)

OK486256	<i>Mafa-I*01:31:01:01</i>	<i>Mane-I*01:13</i> (KF012987)
OK486257	<i>Mafa-I*01:43:03:01</i>	<i>Mamu-I*01:06:24:01</i> (LR990787)
		<i>Mane-I*01:03</i> (KF012987)
		<i>Malo-I*01:01</i> (KT214468)
OK486260	<i>Mafa-I*01:43:03:01nov</i>	<i>Mamu-I*01:06:10</i> (AB541977)
		<i>Mane-I*01:03</i> (KF012987)
		<i>Malo-I*01:01</i> (KT214468)
OK486264	<i>Mafa-B*036:05</i>	<i>Mamu-B*036:05</i> (LT623001)
OK486265	<i>Mafa-I*01:70</i>	<i>Mane-I*01:13</i> (KF012987)
OK486268	<i>Mafa-I*01:35:03:01</i>	<i>Mamu-I*01:01:04</i> (EU429640)
		<i>Maas-I*01:02</i> (KF012975)
OK486269	<i>Mafa-I*01:14:08:02</i>	<i>Mamu-I*01:06:05:02</i> (LR743948)
		<i>Mane-I*01:11</i> (LR743948)
		<i>Malo-I*01:01</i> (KT214468)

注：括号内的序列号是 GenBank 登录号

Note: The serial number in brackets is GenBank accession number.

有的，这些地理来源不同但是序列相似性很高的等位基因可能是通过微调其序列来适应病原体的选择压力。

进化树显示所有的 *Mafa-B* 等位基因分为六大类(图 4)。其中橙色这一大类主要包括 *Mafa-B11L* 和 *Mafa-B*001* 谱系，与用 *B6* 基因特异性引物扩增出来的等位基因结果相符；在粉色这一大类中，*Mafa-B*056* 和 *Mafa-I*01* 谱系被聚类在一起，但是两者又处于不同的分支上说明 2 个谱系的序列相似性很高；在蓝色这一大类中包括 *Mafa-B*034* 和 *Mafa-B*007* 谱系，其与用 *B5* 基因特异性引物的扩增结果相符(表 3 和图 4)。在绿色、紫色和黄色这 3 个大类中分别包含了很多谱系的等位基因，再次说明了 *MHC-B* 等位基因的高度复杂性。因此，将不同的 *Mafa-B* 等位基因分配到不同的 *B* 基因位点非常具有挑战性。

2.6 *Mafa-B* 等位基因重组分析

为了鉴定越南食蟹猴群体中存在的重组事件，对 92 条 *Mafa-B* 序列进行了分析。利用 RDP4 软件鉴定到了 2 种可能的重组事件(表 5)。利用

RIP 软件进一步验证，*Mafa-B*105:01:01:01nov2* 被鉴定为潜在的重组基因(图 5)。*Mafa-B*105:01:01:01nov2* 和 *Mafa-B*156:01:01:01nov* 的外显子 2 序列非常相似，只有 1 个碱基差异；但这两个等位基因的外显子 3 序列差别很大，有 29 个碱基不匹配(图 6)。*Mafa-B*105:01:01:01nov2* 和 *Mafa-B*056:01:01:01* 的外显子 3 序列有 7 个碱基的差异，序列高度相似(图 6)。由于未获得等位基因的内含子序列，所以无法确定准确的断点。此外，对食蟹猴和恒河猴 *B*105*、*B*056* 和 *B*156* 谱系所有的等位基因序列进行系统发育分析。外显子 2 序列的系统发育图显示，*Mafa-B*156* 与 *Mafa-B*105* 的聚类比 *Mafa-B*056* 更接近(图 7A)。外显子 3 序列的系统发育图显示，*Mafa-B*056* 与 *Mafa-B*105* 的聚类比 *Mafa-B*156* 更接近(图 7B)。分析表明，*B*156* 和 *B*105* 谱系在进化树中不区分物种，表明了这 2 个谱系在 2 个猕猴物种之间的系统发育一致性。因此推测，*Mafa-B*105* 可能是由 *Mafa-B*156* 和 *Mafa-B*056* 重组产生的。

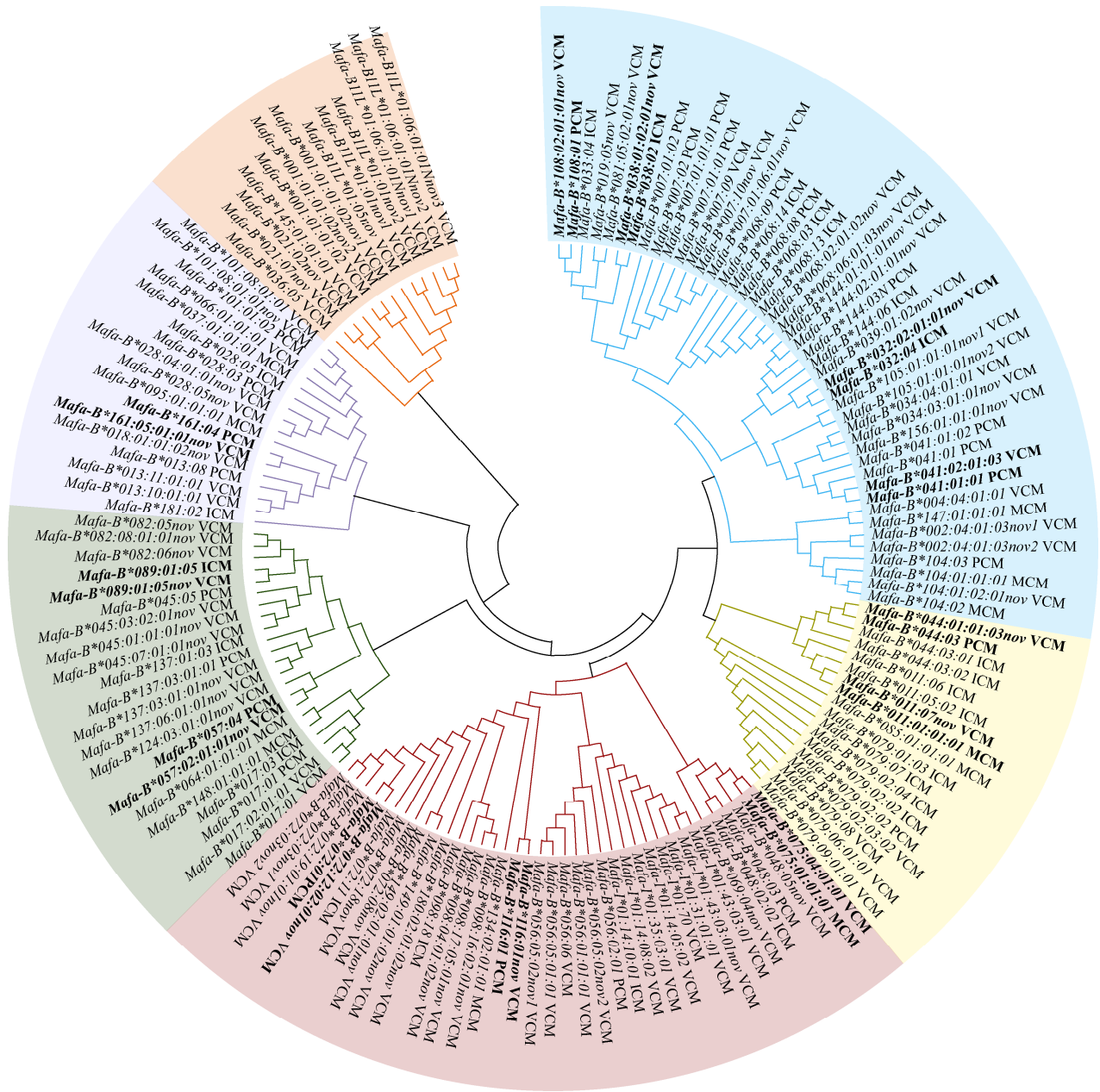


图 4 不同地理来源食蟹猴种群 *Mafa-B* 等位基因之间的系统发育关系 VCM: 越南食蟹猴; ICM: 印度尼西亚食蟹猴; PCM: 菲律宾食蟹猴; MCM: 毛里求斯食蟹猴; 粗体标记的 12 对 *Mafa-B* 等位基因表示其中一个序列在越南种群中发现, 另外一个序列在越南以外的种群中发现

Figure 4 Phylogenetic relationships between *Mafa-B* alleles in *Mafa* populations of different geographical origins. VCM: Vietnamese cynomolgus macaque; ICM: Indonesian cynomolgus macaque; PCM: Philippine cynomolgus macaque; MCM: Mauritian cynomolgus macaque; The 12 pairs of *Mafa-B* alleles marked in bold indicate that one sequence is found in the Vietnamese population and the other sequence is from a population outside Vietnam.

表 5 RDP4 重组分析

Table 5 Recombination test using RDP4

Recombination information	<i>Mafa-B*105:01:01:01nov2</i>	<i>Mafa-I*01:43:03:01nov</i>
Breakpoint begin positions	315	289
Breakpoint end positions	518*	546*
Major parent	<i>Mafa-B*056:01:01:01</i>	<i>Mafa-B*072:18nov</i>
Minor parent	<i>Mafa-B*156:01:01:01nov</i>	<i>Mafa-B*034:04:01:01</i>
Methods	G M S T	R G M C S T

注: *代表实际的断点位置是不确定的; R: RDP; G: GENECONV; M: MaxChi; C: Chimaera; S: SiScan; T: 3Seq
Note: * means the actual breakpoint position is undetermined. R: RDP; G: GENECONV; M: MaxChi; C: Chimaera; S: SiScan; T: 3Seq.

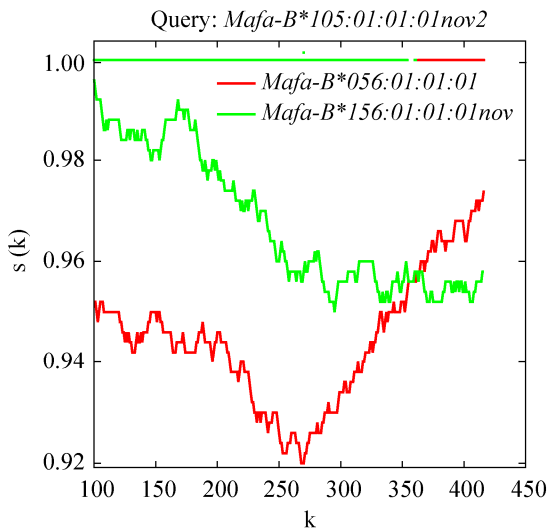


图 5 RIP 重组分析 重组体: *Mafa-B*105:01:01:01nov2*; x 轴(k)表示查询序列在移动窗口中心的位置; y 轴 $s(k)$ 表示序列窗口和每个背景序列之间的相似性

Figure 5 Recombination analysis using RIP. Recombinant: *Mafa-B*105:01:01:01nov2*; The x -axis (k) represents the query sequence position at the center of the moving window; The y -axis, $s(k)$, shows the similarity between that window of sequence and each of the background sequences.

3 讨论与结论

由于猕猴 MHC I 类区域的收缩和扩展, 一个单倍型可能携带不同数量高度相关的 B 基因, 导致很难将等位基因分配给一个特定的 B 基因或位点。因此, 除了 $MHC-I$ 和假基因 $B11L$ 、 $-B12$ 、 $-B16$ 和 $-B17$, $MHC-B$ 基因至今未

引入基因座号。研究表明 $B11L$ 是一个假基因^[10], 而在本研究中采用 $Mafa-B6$ ($B11L$) 基因特异性引物 PCR 扩增出了 $B11L$ 基因的 2 和 3 外显子序列(7/33, 21.2%)。猕猴 B 基因可能在不同的白细胞亚群中表现出不同的表达, 这可能对猕猴物种中的 MHC I 类限制性 T 细胞反应有影响^[36]。通过对 92 个 $Mafa-B$ 序列进行分析, 2 个等位基因经鉴定可能是抗 SIV 感染的保护性同源基因, 1 个等位基因经鉴定可能是 CIA 的保护性同源基因, 这些发现对于在 AIDS 和 CIA 医学研究中使用携带这些保护性同源基因的食蟹猴具有重要意义。 $Mafa-B*007:01$ 被报道在越南来源食蟹猴中是高频等位基因^[25], 在本研究中 $Mafa-B*007:01$ 分布频率较高(5/33, 15.2%), 与之前报道一致。

研究表明, 大多数 $MHC-B$ 等位基因是特定地域的食蟹猴所特有的, $MHC-B$ 等位基因在不同种群中的分布频率也不同, 但是在不同的食蟹猴群体中却仍然观察到了较多的相同等位基因^[18-19,25], 这表明不同地理来源的食蟹猴群体中相同的等位基因可能行使重要的免疫功能以应对环境病原体, 而在不同的猕猴物种中相同的等位基因可能在呈递抗原肽和淋巴细胞识别特定病原体等方面发挥重要作用。遗传学研究表明, 由于食蟹猴在东南亚的地理分布与克拉地峡以北的恒河猴重合导致发生了种间杂交, 一些重要疾病相关基因从恒河猴渗入到食蟹猴中^[37-39]。在本研究中, 食蟹猴与恒河猴中相同

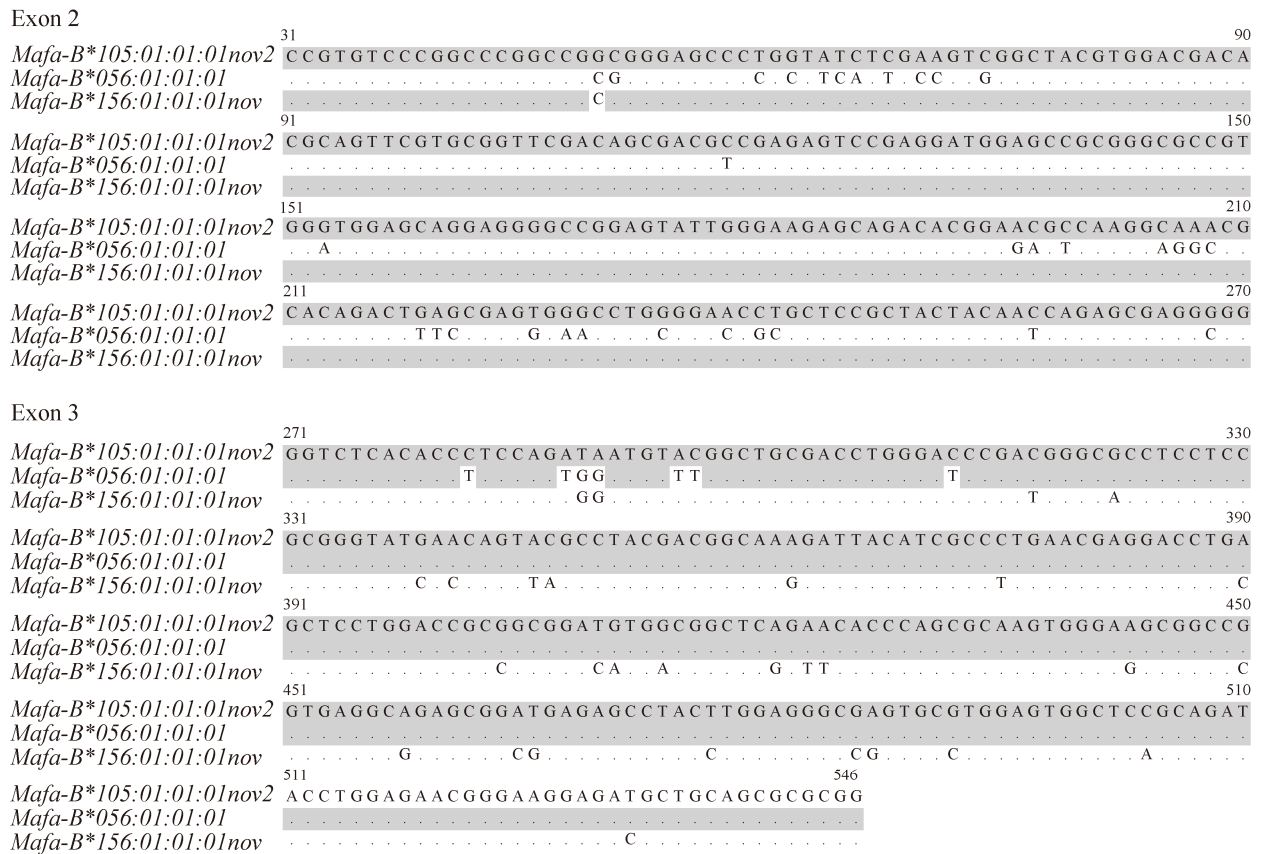


图 6 *Mafa-B*105:01:01:01nov2*、*Mafa-B*156:01:01:01nov* 和 *Mafa-B*056:01:01:01* 等位基因的 2 和 3 外显子序列比对 相同的碱基用“•”表示

Figure 6 The sequences alignment of exons 2 and 3 of *Mafa-B*105:01:01:01nov2*, *Mafa-B*156:01:01:01nov* and *Mafa-B*056:01:01:01*. Identical bases are denoted by “•”.

的等位基因有 28 个，而且相同的等位基因数量比其他猕猴物种多，说明食蟹猴是一种理想的实验替代动物。相同的等位基因数量的增加可能会大大减少研究非人灵长类动物的细胞免疫反应所需的动物总数。

共表达的 MHC I 类等位基因在食蟹猴群体中很常见，作为疾病风险的标志物似乎比单个等位基因更重要。例如，*Mafa-B*430101-B*440101-B*460101* 组合在 52% 的毛里求斯食蟹猴中被鉴定到^[33]。*Mafa-A*25-Mafa-A*29* 组合在 88% 的毛里求斯食蟹猴中存在，这些个体会出现 3 种 SIV 特异性 CD8⁺T 细胞反应^[40]。*Mafa-B*028-Mafa-B*021* 组合在不同地域的食蟹猴及恒

河猴和豚尾猴中均存在^[12-13,17-18,41]。在本研究中，*Mafa-I*01-Mafa-B*034* 组合出现在 6 个个体中，这个组合在柬埔寨食蟹猴中也存在^[17]。此外，*Mafa-I*01-Mafa-B*056* 组合在 9 个印度尼西亚食蟹猴的单倍型中被检测到^[18]，本研究在 3 个个体中鉴定到此组合。这些共表达的等位基因可能在定向筛查或开发新疾病疗法的生物标志物方面具有临床价值。

重组是 MHC I 类等位基因多样化的重要原因之一^[19]。据报道，*Mafa-B*099* 是由 *Mafa-B*054* 和 *Mafa-B*095* 通过重组产生^[42]。为了进一步阐明 *Mafa-B* 等位基因的多样性，我们对重组事件进行了分析。在本研究中，我们鉴定了

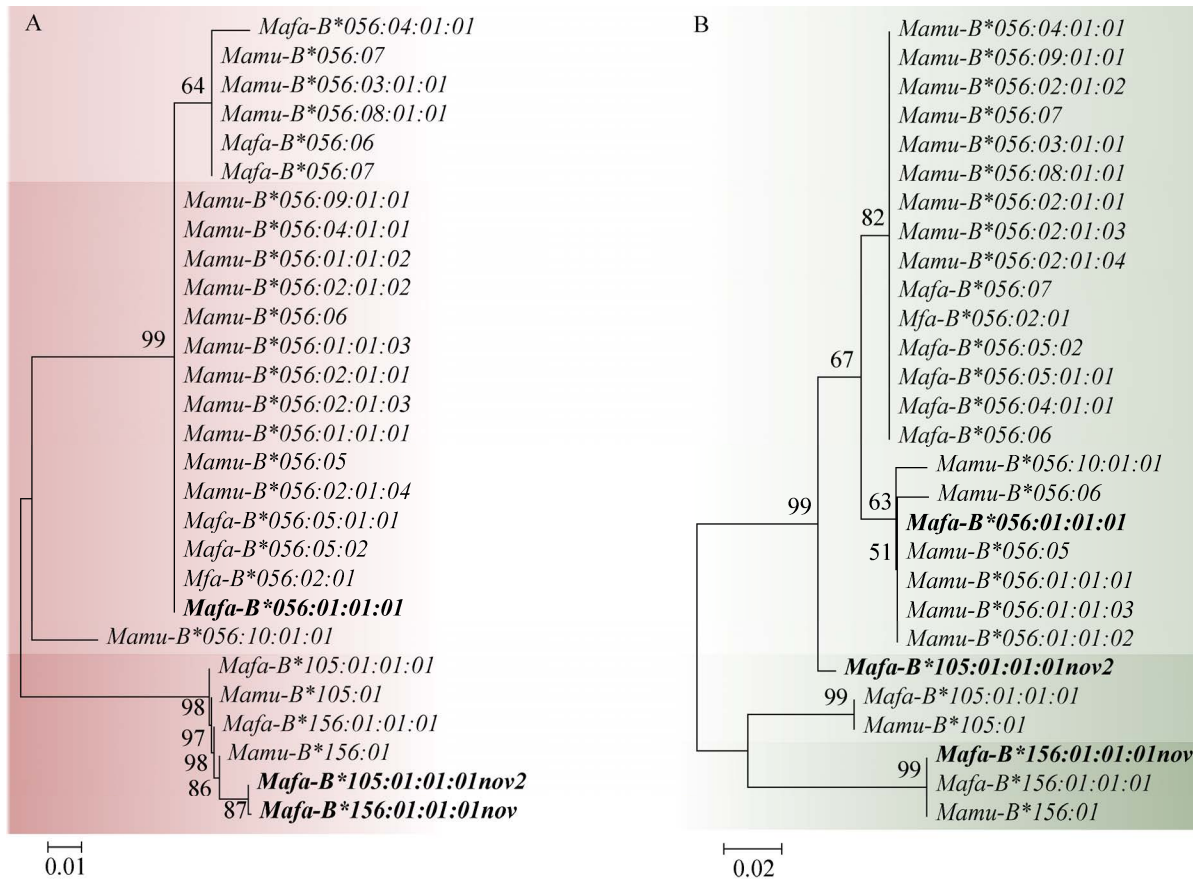


图7 食蟹猴和恒河猴 *B*105*、*B*056* 和 *B*156* 等位基因系统发育分析 A: 外显子 2; B: 外显子 3。分支处数值表示 bootstrap 值; 0.01 和 0.02 代表序列偏差值

Figure 7 Phylogenetic analysis of *B*105*, *B*056* and *B*156* alleles from *Mafa* and *Mamu*. A: Exon 2; B: Exon 3. The numbers at branch nodes present bootstrap values; 0.01 and 0.02 represent the sequence deviation.

一个潜在的重组事件, *Mafa-B*105* 可能是由 *Mafa-B*056* 和 *Mafa-B*156* 通过重组产生的。这些数据表明基因重组在猕猴 *MHC* I 类基因的多样性中发挥重要作用。

不同地域的食蟹猴被用于移植研究^[43-44]。研究者们利用印度尼西亚食蟹猴研究 T 细胞和异体抗体的反应^[45]。菲律宾食蟹猴被用于神经细胞移植研究, 当供体和受体的 *MHC* 基因型匹配时, 移植的结果会有很大改善; 但是当供体和受体的 *MHC* 基因型不匹配时会发生炎症反应^[46]。毛里求斯食蟹猴的 *MHC* 多样性是有限的^[39], 而这种有限的遗传多样性在移植^[47]和

SIV 研究^[48]中具有极大的价值。当食蟹猴被用于医学实验时, 需要慎重考虑不同地域的 *MHC* 基因多样性的信息及它们与各种疾病易感性的关系^[47]。越南食蟹猴的遗传背景是有限的, 本研究为越南来源食蟹猴有限的 *MHC* 免疫遗传信息提供了重要补充, 从而大大提高了该物种作为生物医学研究动物的价值。

本研究将在提供的越南来源食蟹猴 *Mafa-B* 等位基因的基础上, 进一步通过对该物种中 *Mafa-B* 基因遗传进化和功能的研究, 全面了解其免疫功能的复杂性, 阐明 *MHC-B* 基因转录水平对疾病发展的潜在影响。

REFERENCES

- [1] Dijkman K, Vervenne RAW, Sombroek CC, Boot C, Hofman SO, Van Meijgaarden KE, Ottenhoff THM, Kocken CHM, Haanstra KG, Vierboom MPM, Verreck FAW. Disparate tuberculosis disease development in macaque species is associated with innate immunity[J]. *Frontiers in Immunology*, 2019, 10: 2479
- [2] Almond N, Berry N, Stebbings R, Preston M, Ham C, Page M, Ferguson D, Rose N, Li B, Mee ET, Hassall M, Stahl-Hennig C, Athanasopoulos T, Papagatsias T, Herath S, Benlahrech A, Dickson G, Meiser A, Patterson S. Vaccination of macaques with DNA followed by adenoviral vectors encoding simian immunodeficiency virus (SIV) gag alone delays infection by repeated mucosal challenge with SIV[J]. *Journal of Virology*, 2019, 93(21): e00606-19
- [3] Kwon Y, Lee KW, Park H, Son JK, Lee J, Hong J, Park JB, Kim SJ. Comparative study of human and cynomolgus T-cell depletion with rabbit anti-thymocyte globulin (rATG) treatment-for dose adjustment in a non-human primate kidney transplantation model[J]. *American Journal Transplantation Reserach*, 2019, 11(10): 6422-6432
- [4] Rockx B, Kuiken T, Herfst S, Bestebroer T, Lamers MM, Oude Munnink BB, De Meulder D, Van Amerongen G, Van Den Brand J, Okba NMA, et al. Comparative pathogenesis of COVID-19, MERS, and SARS in a nonhuman primate model[J]. *Science*, 2020, 368(6494): 1012-1015
- [5] Francke U, Pellegrino MA. Assignment of the major histocompatibility complex to a region of the short arm of human chromosome 6[J]. *PNAS*, 1977, 74(3): 1147-1151
- [6] Dawkins R, Leelayuwat C, Gaudieri S, Tay G, Hui J, Cattley S, Martinez P, Kulski J. Genomics of the major histocompatibility complex: haplotypes, duplication, retroviruses and disease[J]. *Immunological Reviews*, 1999, 167: 275-304
- [7] Shiina T, Ota M, Shimizu S, Katsuyama Y, Hashimoto N, Takasu M, Anzai T, Kulski JK, Kikkawa E, Naruse T, et al. Rapid evolution of major histocompatibility complex class I genes in primates generates new disease alleles in humans via hitchhiking diversity[J]. *Genetics*, 2006, 173(3): 1555-1570
- [8] Robinson J, Guethlein LA, Cereb N, Yang SY, Norman PJ, Marsh SGE, Parham P. Distinguishing functional polymorphism from random variation in the sequences of >10,000 *HLA-A*, *-B* and *-C* alleles[J]. *PLoS Genetics*, 2017, 13(6): e1006862
- [9] De Groot NG, Otting N, Maccari G, Robinson J, Hammond JA, Blancher A, Lafont BAP, Guethlein LA, Wroblewski EE, Marsh SGE, et al. Nomenclature report 2019: major histocompatibility complex genes and alleles of Great and Small Ape and Old and New World monkey species[J]. *Immunogenetics*, 2020, 72(1-2): 25-36
- [10] Heijmans CMC, De Groot NG, Bontrop RE. Comparative genetics of the major histocompatibility complex in humans and nonhuman primates[J]. *International Journal of Immunogenetics*, 2020, 47(3): 243-260
- [11] Boyson JE, Shufflebotham C, Cadavid LF, Urvater JA, Knapp LA, Hughes AL, Watkins DI. The MHC class I genes of the rhesus monkey. Different evolutionary histories of MHC class I and II genes in primates[J]. *The Journal of Immunology*, 1996, 156(12): 4656-4665
- [12] Doxiadis GGM, De Groot N, Otting N, De Vos-Rouweler AJM, Bolijn MJ, Heijmans CMC, De Groot NG, Van Der Wiel MKH, Remarque EJ, Vangenot C, et al. Haplotype diversity generated by ancient recombination-like events in the MHC of Indian rhesus macaques[J]. *Immunogenetics*, 2013, 65(8): 569-584
- [13] Karl JA, Bohn PS, Wiseman RW, Nimityongskul FA, Lank SM, Starrett GJ, O'Connor DH. Major histocompatibility complex class I haplotype diversity in Chinese rhesus macaques[J]. *G3 (Bethesda)*, 2013, 3(7): 1195-1201
- [14] Daza-Vamenta R, Glusman G, Rowen L, Guthrie B, Geraghty DE. Genetic divergence of the rhesus macaque major histocompatibility complex[J]. *Genome Research*, 2004, 14(8): 1501-1515
- [15] Otting N, Doxiadis GGM, Bontrop RE. Definition of *Mafa-A* and *-B* haplotypes in pedigreed cynomolgus macaques (*Macaca fascicularis*)[J]. *Immunogenetics*, 2009, 61(11): 745-753
- [16] Otting N, De Groot N, De Vos-Rouweler AJM, Louwense A, Doxiadis GGM, Bontrop RE. Multilocus definition of MHC haplotypes in pedigreed cynomolgus macaques (*Macaca fascicularis*)[J]. *Immunogenetics*, 2012, 64(10): 755-765
- [17] Karl JA, Graham ME, Wiseman RW, Heimbruch KE, Gieger SM, Doxiadis GGM, Bontrop RE, O'Connor DH. Major histocompatibility complex haplotyping and long-amplicon allele discovery in cynomolgus macaques from Chinese breeding facilities[J]. *Immunogenetics*, 2017, 69(4): 211-229
- [18] Shortreed CG, Wiseman RW, Karl JA, Bussan HE, Baker DA, Prall TM, Haj AK, Moreno GK, Penedo MCT, O'Connor DH. Characterization of 100 extended major

- histocompatibility complex haplotypes in Indonesian cynomolgus macaques[J]. Immunogenetics, 2020, 72(4): 225-239
- [19] De Groot NG, De Groot N, De Vos-Rouweler AJM, Louwse A, Bruijnesteijn J, Bontrop RE. Dynamic evolution of *Mhc* haplotypes in cynomolgus macaques of different geographic origins[J]. Immunogenetics, 2022, 74(4): 409-429
- [20] Bakker NP, Van Erck MG, Otting N, Lardy NM, Noort RC, 't Hart BA, Jonker M, Bontrop RE. Resistance to collagen-induced arthritis in a nonhuman primate species maps to the major histocompatibility complex class I region[J]. Journal of Experimental Medicine, 1992, 175(4): 933-7
- [21] Nomura T, Matano T. Association of MHC-I genotypes with disease progression in HIV/SIV infections[J]. Frontiers in Microbiology, 2012, 3: 234
- [22] Martin MP, Carrington M. Immunogenetics of HIV disease[J]. Immunological Reviews, 2013, 254(1): 245-264
- [23] Loffredo JT, Maxwell J, Qi Y, Glidden CE, Borchardt GJ, Soma T, Bean AT, Beal DR, Wilson NA, Rehrauer WM, et al. *Mamu-B*08*-positive macaques control simian immunodeficiency virus replication[J]. Journal of Virology, 2007, 81(16): 8827-8832
- [24] Yant LJ, Friedrich TC, Johnson RC, May GE, Maness NJ, Enz AM, Lifson JD, O'Connor DH, Carrington M, Watkins DI. The high-frequency major histocompatibility complex class I allele *Mamu-B*17* is associated with control of simian immunodeficiency virus SIVmac239 replication[J]. Journal of Virology, 2006, 80(10): 5074-5077
- [25] Huang ST, Huang X, Li S, Zhu MJ, Zhuo M. MHC class I allele diversity in cynomolgus macaques of Vietnamese origin[J]. PeerJ, 2019, 7: e7941
- [26] Dzuris JL, Sidney J, Appella E, Chesnut RW, Watkins DI, Sette A. Conserved MHC class I peptide binding motif between humans and rhesus macaques[J]. The Journal of Immunology, 2000, 164(1): 283-291
- [27] Loffredo JT, Sidney J, Bean AT, Beal DR, Bardet W, Wahl A, Hawkins OE, Piaskowski S, Wilson NA, Hildebrand WH, et al. Two MHC class I molecules associated with elite control of immunodeficiency virus replication, *Mamu-B*08* and HLA-B*2705, bind peptides with sequence similarity[J]. The Journal of Immunology, 2009, 182(12): 7763-7775
- [28] Walter L, Ansari AA. MHC and KIR polymorphisms in rhesus macaque SIV infection[J]. Frontiers in Immunology, 2015, 6: 540
- [29] Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA, Charter K, Bib MJ, Bipp M, Keiser T, Butner M. Practical *Streptomyces* Genetics[M]. 2nd ed. Norwich: The John Innes Foundation, 2000
- [30] He QD, Chen M, Lin XG, Chen ZG. Allele-specific PCR with a novel data processing method based on difference value for single nucleotide polymorphism genotyping of *ALDH2* gene[J]. Talanta, 2020, 220: 121432
- [31] Mothé BR, Weinfurter J, Wang CX, Rehrauer W, Wilson N, Allen TM, Allison DB, Watkins DI. Expression of the major histocompatibility complex class I molecule *Mamu-A*01* is associated with control of simian immunodeficiency virus SIVmac239 replication[J]. Journal of Virology, 2003, 77(4): 2736-2740
- [32] Maness NJ, Walsh AD, Rudersdorf RA, Erickson PA, Piaskowski SM, Wilson NA, Watkins DI. Chinese origin rhesus macaque major histocompatibility complex class I molecules promiscuously present epitopes from SIV associated with molecules of Indian origin; implications for immunodominance and viral escape[J]. Immunogenetics, 2011, 63(9): 587-597
- [33] Krebs KC, Jin ZY, Rudersdorf R, Hughes AL, O'Connor DH. Unusually high frequency MHC class I alleles in Mauritian origin cynomolgus macaques[J]. The Journal of Immunology, 2005, 175(8): 5230-5239
- [34] Shiina T, Yamada Y, Aarnink A, Suzuki S, Masuya A, Ito S, Ido D, Yamanaka H, Iwatani C, Tsuchiya H, Ishigaki H, Itoh Y, Ogasawara K, Kulski JK, Blancher A. Discovery of novel *MHC-class I* alleles and haplotypes in Filipino cynomolgus macaques (*Macaca fascicularis*) by pyrosequencing and Sanger sequencing[J]. Immunogenetics, 2015, 67(10): 563-578
- [35] Campbell KJ, Detmer AM, Karl JA, Wiseman RW, Blasky AJ, Hughes AL, Bimber BN, O'Connor SL, O'Connor DH. Characterization of 47 MHC class I sequences in Filipino cynomolgus macaques[J]. Immunogenetics, 2009, 61(3): 177-187
- [36] Greene JM, Wiseman RW, Lank SM, Bimber BN, Karl JA, Burwitz BJ, Lhost JJ, Hawkins OE, Kunstman KJ, Broman KW, et al. Differential MHC class I expression in distinct leukocyte subsets[J]. BMC Immunology, 2011, 12: 39
- [37] Yan GM, Zhang GJ, Fang XD, Zhang YF, Li C, Ling F, Cooper DN, Li QY, Li Y, Van Gool AJ, et al. Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques[J]. Nature Biotechnology, 2011, 29(11): 1019-1023
- [38] Fan ZX, Zhou AB, Osada N, Yu JQ, Jiang J, Li P, Du

- LM, Niu LL, Deng JB, Xu HL, et al. Ancient hybridization and admixture in macaques (genus *Macaca*) inferred from whole genome sequences[J]. *Molecular Phylogenetics and Evolution*, 2018, 127: 376-386
- [39] Osada N, Hettiarachchi N, Adeyemi Babarinde I, Saitou N, Blancher A. Whole-genome sequencing of six Mauritian cynomolgus macaques (*Macaca fascicularis*) reveals a genome-wide pattern of polymorphisms under extreme population bottleneck[J]. *Genome Biology and Evolution*, 2015, 7(3): 821-830
- [40] Burwitz BJ, Pendley CJ, Greene JM, Detmer AM, Lhost JJ, Karl JA, Piaskowski SM, Rudersdorf RA, Wallace LT, Bimber BN, et al. Mauritian cynomolgus macaques share two exceptionally common major histocompatibility complex class I alleles that restrict simian immunodeficiency virus-specific CD8⁺ T cells[J]. *Journal of Virology*, 2009, 83(12): 6011-6019
- [41] Semler MR, Wiseman RW, Karl JA, Graham ME, Gieger SM, O'Connor DH. Novel full-length major histocompatibility complex class I allele discovery and haplotype definition in pig-tailed macaques[J]. *Immunogenetics*, 2018, 70(6): 381-399
- [42] Orysiuk D, Lawrence J, Prashar T, Spangelo L, Pilon R, Fournier J, Rud E, Sandstrom P, Plummer FA, Luo M. Evidence of recombination producing allelic diversity in MHC class I *Mafa-B* and *-A* alleles in cynomolgus macaques[J]. *Tissue Antigens*, 2012, 79(5): 351-358
- [43] Kato Y, Griesemer AD, Wu A, Sondermeijer HP, Weiner JI, Duran-Struuck R, Martinez M, Slate AR, Romanov A, Lefkowitz JH, et al. Novel H-shunt venovenous bypass for liver transplantation in cynomolgus macaques[J]. *Comparative Medicine*, 2017, 67(5): 436-441
- [44] Matsunami M, Rosales IA, Adam BA, Oura T, Mengel M, Smith RN, Lee H, Cosimi AB, Colvin RB, Kawai T. Long-term kinetics of intragraft gene signatures in renal allograft tolerance induced by transient mixed chimerism[J]. *Transplantation*, 2019, 103(11): e334-e344
- [45] Ezzelarab MB, Zhang H, Guo H, Lu L, Zahorchak AF, Wiseman RW, Nalesnik MA, Bhama JK, Cooper DKC, Thomson AW. Regulatory T cell infusion can enhance memory T cell and alloantibody responses in lymphodepleted nonhuman primate heart allograft recipients[J]. *American Journal of Transplantation*, 2016, 16(7): 1999-2015
- [46] Morizane A, Kikuchi T, Hayashi T, Mizuma H, Takara S, Doi H, Mawatari A, Glasser MF, Shiina T, Ishigaki H, et al. MHC matching improves engraftment of iPSC-derived neurons in non-human primates[J]. *Nature Communications*, 2017, 8: 385
- [47] Shiina T, Blancher A. The cynomolgus macaque MHC polymorphism in experimental medicine[J]. *Cells*, 2019, 8(9): 978
- [48] Wiseman RW, Wojcechowskyj JA, Greene JM, Blasky AJ, Gopon T, Soma T, Friedrich TC, O'Connor SL, O'Connor DH. Simian immunodeficiency virus SIVmac239 infection of major histocompatibility complex-identical cynomolgus macaques from Mauritius[J]. *Journal of Virology*, 2007, 81(1): 349-361