



恒河猴 *samhd1* 基因编码区多态性及其对 蛋白结构和功能的影响

熊圣文, 王 卫, 董志会, 陈 霆, 魏 强

(北京协和医学院比较医学中心, 中国医学科学院医学实验动物研究所, 卫生部人类疾病比较医学重点实验室,
国家中医药管理局人类疾病动物模型三级实验室, 北京 100021)

【摘要】 目的 筛查恒河猴 *samhd1* 基因编码区单核苷酸多态性 (coding-region single nucleotide polymorphism, cSNP), 研究其对蛋白结构和功能的影响。方法 提取恒河猴外周血 RNA, RT-PCR 扩增, 测序比对, 确定基因编码区变异位点; 通过蛋白质结构分析软件对该蛋白结构进行分析; 使用 SNaPshot 方法筛查 138 只恒河猴基因背景。结果 序列比对发现 5 个非同义突变位点 C15G (2.17%)、C320T (6.88%)、G547A (13.41%)、A552T (4.35%)、T839C (17.39%); 其中除 C15G 外的其它四个 cSNP 位点均处于 SAMHD1 蛋白重要结构域上, 且 G547A、A552T、T839C 三个 cSNP 位点改变 SAMHD1 蛋白二级结构, 可能影响 SAMHD1 蛋白功能。结论 恒河猴 *samhd1* 基因编码区存在三个可能影响 SAMHD1 蛋白功能的 cSNP 位点, 为后续研究提供了参考信息。

【关键词】 SAMHD1; cSNP; 蛋白质结构; 功能; 恒河猴

【中图分类号】 R33 **【文献标识码】** A **【文章编号】** 1671-7856(2014) 04-0001-04

doi: 10.3969/j.issn.1671.7856.2014.004.001

Polymorphisms of coding region in *samhd1* gene of rhesus macaques and their effects on protein structure and function

XIONG Sheng-wen, WANG Wei, DONG Zhi-hui, CHENG Ting, WEI Qiang

(Comparative Medicine Center, Peking Union Medical College (PUMC) & Institute of Medical Laboratory Animal Science, Chinese Academy of Medical Sciences (CAMS); Key Laboratory of Human Diseases Comparative Medicine, Ministry of Health; Key Laboratory of Human Diseases Animal Models, State administration of Traditional Chinese Medicine, Beijing 100021, China)

【Abstract】 **Objective** To study the impact of cSNP on the structure and function of *samhd1* gene in rhesus macaques. **Methods** RNA was extracted from peripheral blood of rhesus macaques. The cSNP sites were detected and confirmed through reverse transcription PCR, sequencing and sequence alignment. Protein structure and function were analyzed with Predictprotein and SWISS-MODEL. Finally, the functional cSNP sites were screened by SNaPshot technique in a larger sample scale. **Results** Five non-synonymous mutation sites (C15G (2.17%), C320T (6.88%), G547A (13.41%), A552T (4.35%) and T839C (17.39%)) were found, respectively. Four of them, except C15G, were included in the key structure domains of SAMHD1. Meanwhile, G547A, A552T, T839C had secondary structure changed and were likely to impact the function of SAMHD1. **Conclusion** Some important cSNP sites are found in *samhd1* gene of

【基金项目】 国家科技重大专项课题(2012ZX10004501-001, 2013ZX10004-608)。

【作者简介】 熊圣文(1989-), 男, 硕士研究生, 比较医学专业。

【通讯作者】 魏强, 教授, 博士生导师, 研究方向: 实验动物病毒学。Email: weiqiang0430@sohu.com。

rhesus macaques, which may impact on the function of SAMHD1.

【Key words】 SAMHD1; cSNP; Structure; Function; Rhesus macaque

SAMHD1 (SAM domain and HD domain 1) 是近年来发现的对 HIV-1 等逆转录病毒复制具有明显限制作用的天然抗病毒蛋白之一^[1,2],而在恒河猴体内也存在抑制猴免疫缺陷病毒 (simian immunodeficiency virus, SIV) 复制的 *samhd1* 基因^[3]。前期研究表明, Trim5 α 和 Apobec3G 等天然抗病毒蛋白在进化过程中,会出现影响抗病毒功能的核苷酸突变;进化分析也证实, *samhd1* 基因部分氨基酸残基位点可能影响 Vpx 对 SAMHD1 的敏感性^[3]。但有关恒河猴 *samhd1* 基因多态性,及对蛋白结构和功能的影响尚未见报道。本研究旨在通过基因测序,序列比对,蛋白质结构预测等技术找到恒河猴 *Samhd1* 基因中影响抗 SIV 功能的 cSNP 位点,从而为探索 SAMHD1 抗病毒机制和优化 SIV/SHIV/恒河猴模型提供研究基础。

1 材料和方法

1.1 实验动物与样品采集

中国恒河猴 138 只,购自北京协尔鑫生物资源研究所(合格证编号:SCXK(京)2005-2005);分别采集动物外周血 1 mL, EDTA 抗凝,样品采集后 4 h 内提取 RNA。

1.2 外周血 RNA 的提取及反转录

实验猴 EDTA 抗凝全血使用 RNA blood mini kit (购自 Qiagen 公司)进行总 RNA 的提取,提取过程参照试剂盒说明书;提取 RNA 后立即对其进行反转录。

1.3 RT-PCR 扩增

Samhd1 基因扩增采用巢式 PCR,第一轮 RT-PCR 使用 One Step RNA PCR Kit (AMV) (TaKaRa, 大连)进行,操作过程参照试剂盒说明书;第二轮 PCR 使用 KOD-Plus 试剂盒 (TOYOBO, 上海)进行,操作过程参照试剂盒说明书。*Samhd1* 基因引物由引物设计软件 Primer Premier 5.0 设计完成,由 Invitrogen 公司合成。外套引物序列为:outer-F 5'-CCC GAA GGG CTC AAC TGT CA-3'; outer-R 5'-AGA CTT TAA TAA TAC TAA AAA TAG GCA AGG-3',退火温度为 44.5 $^{\circ}$ C。内套引物序列为:inner-F 5'-CAA TAC TGC TTG GAC TCC CCG CC-3'; inner-R 5'-ATT CTA GGA GTG AAG CCC CTA AAT GAA TT-

3',退火温度为 51 $^{\circ}$ C。PCR 在 System 9700 PCR 仪上进行 (Applied Biosystems 公司)。

1.4 序列比对和生物软件分析

DNA 测序由北京诺赛基因公司完成。利用 DNASTAR 软件进行核苷酸和氨基酸序列比对^[4],标准恒河猴 [*Macaca mulatta*] SAMHD1 cDNA 序列为:NM_001271642.1;标准恒河猴 [*Macaca mulatta*] SAMHD1 氨基酸序列为:NP_001258571。

分别用 Predictprotein、SWISS-MODEL 软件进行蛋白质二级和三级结构的预测。

1.5 SNaPshot 分型检测

SNP 的分型检测采用 SNaPshot 技术,即基于荧光标记单碱基延伸原理的分型技术。该检测过程交由北京阅微基因公司,基于 3730XL 平台完成。

2 结果

2.1 中国恒河猴 *Samhd1* 基因编码区 cSNP 位点的分布

提取 29 只恒河猴全血 RNA,逆转录 PCR 后进行测序比对(参照序列为 NCBI 上公布的 *Samhd1* cDNA 序列, NP_001258571),一共找到 11 个突变位点,分别位于基因的 N 端、SAM 结构域、链接区、HD 结构域和 C 端(表 1);通过使用 DNASTAR 软件进行翻译和氨基酸序列比对,发现其中 5 个突变位点为非同义突变,在 *Samhd1* 编码区上的位置分别是:15 bp (C/G)、320 bp (C/T)、547 bp (G/A)、552 bp (A/T)、839 bp (T/C)^[5](图 1)。其中 G183S、E184D、V280A 分布在 HD 结构域上;S107L 分布在 SAM 结构域上;D5E 分布在 N 端非结构域上。

2.2 非同义 cSNP 对 SAMHD1 蛋白二级结构的影响

Predictprotein 软件分析结果显示,5 个非同义 cSNP 位点突变影响了 SAMHD1 蛋白的二级结构,具体的,在 264aa 位置附近一个 α 螺旋结构变成 β 片层,530aa 附近增加了一个 α 螺旋。295aa 附近的蛋白质结合位点移动到 540aa 附近,其它结构均未发生变化(图 2,见彩插 1)。但当单独出现 280aa 或 183aa 和 184aa 的突变时,软件分析结果显示并不引起此二级结构变化(图未列出)。



图 1 恒河猴 SAMHD1 的蛋白质结构及 SNP 分布情况

Fig. 1 Protein structure and SNP distribution in SAMHD1 of the rhesus macaques

表 1 恒河猴 *samhd1* 基因突变位点
Tab. 1 *samhd1* nucleotide mutants of the rhesus macaques

结构域	突变位点(bp)	核苷酸突变	氨基酸位置(aa)	氨基酸突变
N 端	15	C -> G	5	D-E(天冬氨酸-谷氨酸)
SAM 区	168	C -> T	56	/
	320	C -> T	107	S-L(丝氨酸-亮氨酸)
链接区	396	C -> T	132	/
	528	G -> A	176	/
HD 区	547	G -> A	183	G-S(甘氨酸-丝氨酸)
	552	A -> T	184	E-D(谷氨酸-天冬氨酸)
	839	T -> C	280	V-A(缬氨酸-丙氨酸)
	1125	A -> G	375	/
C 端	1692	C -> G	564	/
	1752	C -> T	584	/

2.3 非同义 cSNP 对 SAMHD1 蛋白功能的影响

为预测 cSNP 对 SAMHD1 蛋白功能的影响,使用 SWISS-MODEL 软件对 SAMHD1 蛋白不同结构域进行结构预测。

图 3A 所示 *samhd1* 基因 SAM 结构域的三级结构分析图,绿色标出文献中通过进化分析找到的影响 Vpx 对 SAMHD1 敏感性的位点,其中,包括红圈中的 S107L 位点,说明 S107L 影响病毒 Vpx 蛋白对 SAMHD1 的敏感性。

图 3B 为 *samhd1* 基因 HD 结构域的三级结构分析图,可以看到 G183S、E184D、V280A 三个位点均分布在蛋白质表面,且和文献中已知的重要位点(绿色标出)距离较远,其中 V280A 处于蛋白质活性中心(凹槽部位)的袋口处,类似于盒盖的结构上,可能会影响到 SAMHD1 的脱氧核苷酸三磷酸水解酶催化活性;而 G183S、E184D 两个位点均分布在凹槽的背面,但文献中报道的一些可能影响 Vpx 对 SAMHD1 敏感性的位点也位于类似位置(如图 3B “凹槽”下方的绿色位点),所以不能排除 G183S、E184D 两个位点会影响 SAMHD1 与 Vpx 结合的可能性(图 3 见彩插 1)。

2.4 基因分型及 cSNP 位点的频率测定

根据发现的核苷酸有义突变位点,利用 SNaPshot 技术对 138 只中国恒河猴进行了基因分型

并验证是否为 cSNP 位点。通过 SNaPshot 检测,发现 138 只中国恒河猴中 C15G 基因频率占 2.17%; C320T 的基因频率为 6.88%; G547A 的基因频率为 13.41%; A552T 的基因频率为 4.35%; T839C 的基因频率为 17.39%,这 5 个位点均可判定为 SNP(基因频率大于 1%)。另外还发现 G547A 和 T839C 存在较强的连锁关系,连锁频率占 21.01%。对基因分型结果进行分析,本实验共找到了 6 种突变基因型(根据 5 个 cSNP 在基因中的前后顺序,将正常基因型表示为:… C…C…GA…T…),详情见表 2。

表 2 恒河猴 *samhd1* 突变基因型及比率
Tab. 2 Mutant genotype and ratio of rhesus macaque *samhd1*

名称	突变基因型 (…15b p…320 bp…547/552 bp…839 bp…)	比率(%)
RM1	… <u>C</u> …C…GA…T…	1.81%
RM2	…C… <u>T</u> …GA…T…	5.80%
RM3	…C…C… <u>AA</u> …T…	5.07%
RM4	…C…C… <u>G</u> <u>T</u> …T…	3.26%
RM5	…C…C…GA… <u>C</u> …	7.97%
RM6	…C…C… <u>AA</u> … <u>C</u> …	6.88%

注:突变核苷酸位点用“_”标出。

Note: Nucleotide mutants are marked with “_”.

3 讨论

SAMHD1 主要在单核细胞、单核细胞来源的巨噬细胞、树突状细胞等髓系细胞以及静息 CD4⁺ T 细胞中表达^[1]。SAMHD1 具有磷酸水解酶活性,受

dGTP 的调控,它对 HIV-1 复制的限制是通过降解细胞内 dNTP 的水平^[6,7],使胞内 dNTPs 的水平低于病毒复制的需要水平,从而切断 HIV-1 逆转录所需底物的来源;这一过程被称为“核苷酸仓库耗竭”。研究发现 SAMHD1 对 HIV-1 的复制具有明显的限制作用,但可以被 HIV-2 和 SIV 编码的 Vpx 蛋白拮抗^[2]。另外,有些 Vpr 蛋白也能起到类似作用^[3]。2011 年, Powell et al 通过功能实验找到了 20 多个与 SAMHD1 蛋白功能有关的氨基酸残基位点^[8]。2012 年, Lim et al 通过正向选择分析找到 6 个氨基酸残基位点可能影响 Vpx 对 SAMHD1 的敏感性^[3]。因此,氨基酸残基位点的突变对于 SAMHD1 的抗病毒功能将可能产生很大影响。

非同义 cSNP 碱基序列的改变可使以其为模板翻译的氨基酸序列发生改变,从而影响蛋白质的功能。这种改变常是导致生物性状改变的直接原因。因此,本研究主要分析了 SAMHD1 分子的编码区非同义 SNP 位点。

和人类相似,在进化过程中恒河猴 SAMHD1 分子上也存在着大量的 SNP 位点。2012 年, Coon et al 在分析人 *samhd1* 基因多态性时,发现 SNP (rs1291142) 与 SAMHD1 表达量显著相关^[7]。因此,我们对恒河猴 SAMHD1 分子的编码区突变情况进行了大样本的筛查,总共有 5 个位点被鉴定为非同义 cSNP,在群体中均占有较高的频率,其中 G183S、E184D、V280A 分布在具磷酸水解酶活性的 HD 结构域上;S107L 分布在参与蛋白-蛋白或蛋白-RNA 相互作用的 SAM 结构域上。

为了判断这些 cSNP 是否能影响到蛋白质的功能,我们先进行了蛋白高级结构和功能预测分析。用软件对 SAMHD1 二级、三级结构的分析发现,5 个 cSNP 中 G183S、E184D、V280A 三个位点最有可能影响蛋白质的功能和结构,在二级结构上,这三个位点所处的 HD 结构域发生了一个 α 螺旋到 β 片层的转变,一个蛋白质结合位点也因为这三个 cSNP 的出现而发生了移位;在三级结构上,V280A 正好处于蛋白质活性中心(凹槽部位)的袋口处,类似于盒盖的结构上,可能会影响到 SAMHD1 的磷酸水解酶催化活性。另外,发现 S107L 与文献中报道的可能影响病毒 Vpx 蛋白对 SAMHD1 敏感性的位点重合,说明此位点也有较大可能会改变蛋白质的功

能。D5E 未分布于重要的蛋白结构域上,目前还没有能改变蛋白功能的证据支持,但也有文献报道,SAMHD1 蛋白的 N 端能够与艾滋病毒相互作用,所以有待进一步的实验验证。如果实验验证这些位点与抗病毒功能有关,则说明该蛋白与 SIV 在相互进化过程中发生了选择作用。

这些分析为之后蛋白功能实验提供了大量的参考依据,同时也为 SAMHD1 抗病毒机制的进一步研究发掘了新的靶点。我们后续还将分别对在 SAMHD1 上发现的 6 种基因型进行蛋白功能实验验证。

参考文献:

- [1] Baldauf HM, Pan X, Erikson E, et al. SAMHD1 restricts HIV-1 infection in resting CD4(+) T cells [J]. Nat Med, 2012, 18(11):1682-1687.
- [2] Laguette N, Sobhian B, Casartelli N, et al. SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx [J]. Nature, 2011. 474(7353):654-657.
- [3] Lim ES, Fregoso OI, McCoy CO, et al. The ability of primate lentiviruses to degrade the monocyte restriction factor SAMHD1 preceded the birth of the viral accessory protein Vpx [J]. Cell Host Microbe, 2012, 11(2):194-204.
- [4] 刘克剑,丛喆,金光,等. 表现神经症状的 SIVmac251 感染猴大脑基底节病毒 gp120 序列变异分析 [J]. 中国实验动物学报, 2012, 2:37-42.
- [5] 李佩璐,陈倩倩,张驰宇. 髓系单核细胞来源的 HIV-1 限制性因子—SAMHD1 [J]. 动物学研究, 2012, 33(5):537-541.
- [6] Powell RD, Holland PJ, Hollis T, et al. Aicardi-Goutieres syndrome gene and HIV-1 restriction factor SAMHD1 is a dGTP-regulated deoxynucleotide triphosphohydrolase [J]. J Biol Chem, 2011, 286(51):43596-43600.
- [7] Goldstone DC, Ennis-Adeniran V, Hedden JJ, et al. HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase [J]. Nature, 2011, 480(7377):379-382.
- [8] Coon S, Wang D, Wu L. Polymorphisms of the SAMHD1 gene are not associated with the infection and natural control of HIV type 1 in Europeans and African-Americans [J]. AIDS Res Human Retroviruses, 2012, 28(12):1565-1573.

[修回日期]2014-02-27