

中国部分地区 HIV 患者 B' 和 B'/C 亚型毒株 *tat* 第一外显子基因变异与 HIV 疾病进展关系的研究

韩晓旭 代娣 张岩 张旻 张子宁 刁莹莹 耿文清 尚红

【摘要】 目的 了解中国 B' 和 B'/C 亚型 HIV/AIDS 患者 *tat* 第一外显子的基因序列及其二级结构的特征和变异特点,探讨其与 HIV-1 感染疾病进展之间的关系。方法 从辽宁、吉林和云南省 HIV-1 感染者中选取病情呈缓慢进展的 B' 亚型感染者 8 例和 B'/C 亚型感染者 5 例,选择年龄、性别感染时间与前二者匹配的病情呈典型进展的 B' 亚型感染者 26 例和 B'/C 亚型感染者 9 例。采集外周静脉血,提取前病毒 DNA,用巢式聚合酶链反应扩增 HIV-1 的 *tat* 基因,纯化后直接进行基因序列测定,序列编辑后翻译成氨基酸序列,进行氨基酸变异情况分析和二级结构预测。结果 B' 和 B'/C 亚型缓慢进展者、典型进展者 Tat 第一外显子中发现多种氨基酸替换,但除 A58T 外均未显示出与病毒载量以及疾病进展的明确相关性。23N、31S、32Y、46F 变异均显示出亚型特异性;Tat 蛋白的二级结构未发现规律性变化。结论 中国 HIV/AIDS 患者 *tat* 第一外显子某些位点的基因变异,如 A58T 可能与病毒载量以及疾病进展有关,Tat 蛋白的二级结构可能与 HIV 感染后的疾病进展无明显关系。

【关键词】 人类免疫缺陷病毒 1 型; *tat* 基因; 变异; 二级结构

Study on the relationship between the polymorphisms and secondary structure of *tat* exon-1 gene and HIV/AIDS progress in subtype B' and B'/C HAN Xiao-xu, DAI Di, ZHANG Yan, ZHANG Min, ZHANG Zi-ning, DIAO Ying-ying, GENG Wen-qing, SHANG Hong. Key Laboratory of Immunology of AIDS, Ministry of Health, the First Affiliated Hospital, China Medical University, Shenyang 110001, China

Corresponding author: SHANG Hong, Email: p3lab@yeah.net

【Abstract】 **Objective** To study the polymorphisms and secondary structure of human immunodeficiency virus (HIV-1) *tat* exon 1 among subtype B' and B'/C HIV-1 infected people in China and to explore the relationship between the polymorphism of *tat* exon 1 and the disease progression. **Methods** 8 subtype B' and 5 B'/C HIV-1 infected patients with slow disease progression were selected from Liaoning, Jilin and Yunnan province. 26 subtype B' and 9 B'/C HIV-1 infected patients with similar sex, age but with typical disease progression were selected. Provirus was extracted from the whole blood. The gene sequences of the Tat exon 1 were amplified by nest-polymerase chain reaction (nest-PCR). Products were purified and sequenced directly. The sequences were aligned, translated, amino acid substitution were analyzed and secondary structures were predicted. **Results** Many amino acid substitution could be found in the exon 1 of Tat in HIV-1 subtype B' and B'/C recombinant strain infected persons with different disease progression except A58T, none of them showed definitely relationship with HIV viral load and disease progression. 23N, 31S, 32Y and 46F were subtype-specific substitutions. No characteristic secondary structure of exon 1 of Tat was found. **Conclusion** Some of the mutations of *tat* exon 1 might be related to HIV viral load and disease progression. However, there was no relationship found between the secondary structure of Tat protein and the disease progression.

【Key words】 Human immunodeficiency virus; *tat* gene; Mutation; Secondary structure

基金项目:国家“十五”科技攻关课题资助项目(2004BA719A12);辽宁省优秀青年科研人才培养资金资助项目(3040011);教育部博士点专项课题资助项目(20040159005)

作者单位:110001 沈阳,中国医科大学附属第一医院卫生部艾滋病免疫学重点实验室

通讯作者:尚红,Email: p3lab@yeah.net

Tat 是 HIV-1 的转录过程中起着关键作用的调节蛋白。Tat 通过转录反应因子与某些细胞因子作用形成蛋白-RNA 复合物,活化转录启动因子,加强 RNA 多聚酶 II 的活性,从而明显促进病毒转录。Tat 蛋白的高水平表达可以促进大量病毒的产生,而 *tat* 缺失或突变则可能使病毒完全不复制。国外研究表明 HIV-1 感染者体内的病毒复制水平决定疾病的进程^[1]。快速进展期患者体内的 HIV-1 复制水平很高,致使 CD4⁺ T 淋巴细胞数量迅速下降进入 AIDS 期^[2]。无症状感染者体内的病毒载量通常不高,病毒复制效率也低^[3]。有效的高水平的 HIV-1 复制是 AIDS 进展的先决条件之一。作为编码 Tat 蛋白的 *tat* 基因,其变异是否会影响改变病毒的复制并影响疾病的进展,国内外相关报道较少。本研究通过氨基酸序列分析,初步研究了我国 B' 和 B'/C 亚型 HIV-1 感染者病毒基因组中 *tat* 区第一外显子氨基酸序列的特征、突变的种类,并进行 Tat 蛋白二级结构特征分析;探讨 HIV-1 *tat* 区第一外显子氨基酸多态性及二级结构变化与 HIV 感染疾病进展的相关性。

资料与方法

1. 病例选择:选取 2000-2005 年辽宁、吉林、云南省未接受抗病毒治疗的 HIV 感染者/AIDS 患者的流行病学资料和临床资料。病情呈典型进展者 34 例,缓慢进展者 13 例(表 1)。缓慢进展组标准为感染期超过 10 年,未接受抗病毒治疗,CD4⁺ 细胞 $\geq 500 \times 10^6/L$ 。采集全血标本共 47 份,EDTA 抗凝 -80℃ 冻存。

2. 核酸提取和扩增:以 QIAamp DNA Blood Mini Kit(QIAgen,德国)试剂盒,从 EDTA 抗凝全血标本提取基因组 DNA, *tat* 区基因扩增的引物及反应条件参见文献[4]。QIAquick Gel Extraction Kit 按操作说明回收目的 DNA 片段。

3. HIV-1 *tat* 基因序列测定:取 50 fmol PCR 扩增片段为模板, *tat*-B 为测序引物, BigDye Termintor Sequencing Kit(ABI,美国)按操作说明应用双脱氧终止法进行测序反应,使用 ABI 3130 全自

动遗传分析仪进行毛细管电泳并读取序列。

4. HIV/AIDS Tat 区第一外显子核苷酸、氨基酸序列及其二级结构特征分析:

(1)参照测序谱图进行基因序列修订后,下载 HIV 序列数据(<http://hiv-web.lanl.gov>)中 HIV-1 各亚型流行株参考序列,用 BIOEDIT 软件包中 CLUSTAL W 程序进行排列比对,在核苷酸序列中人工引入 gaps 以保持翻译完整性。统计测得的 Tat 区第一外显子核苷酸及氨基酸突变位置和种类。

(2)以 NNPREDEICT 程序(<http://ncbi.nlm.nih.gov>)预测我们测得的 Tat 区第一外显子氨基酸序列的二级结构,并进行分组比对。

5. 统计学分析:SPSS 11.5 统计软件进行统计学分析,病毒载量取对数值后,呈正态分布。*t* 检验、 χ^2 检验,以 $P < 0.05$ 为有统计学意义。

结果

1. 临床资料:13 例病情呈缓慢进展者通过 *gag* 和 *env* 基因区的分析确定有 B' 亚型 8 例, B'/C 亚型 5 例。35 例病情呈典型进展者中有 B' 亚型 26 例, B'/C 亚型 9 例。缓慢进展组的外周血 CD4⁺ T 细胞数为 (624 ± 180) 个/ μ l,病毒载量 (4.21 ± 0.75) lg 拷贝/ml,而典型进展组的外周血 CD4⁺ T 细胞数为 (226 ± 145) 个/ μ l,病毒载量 (5.13 ± 0.89) lg 拷贝/ml。缓慢进展组和典型进展组之间病毒载量水平的差异有统计学意义(表 1)。

2. 氨基酸变异:缓慢进展组、典型进展组 Tat 蛋白氨基酸序列具有较高同源性,没有发现移码突变和提前出现的终止密码子,只在 1 例典型进展的 B' 亚型感染者的 Tat 蛋白发现 1 个氨基酸缺失。B' 亚型核苷酸序列翻译成氨基酸后与 B 亚型共享序列比较(图 1),缓慢进展组中发现 K19R、A21S、R57S、A58S、H6N、L69V 各 1 例,典型进展组未发现此变异。典型进展组发现 2 例 Q17L, K19N、T20P 各 4 例, A21L/R、A26 F/C、I39T、A58P、V67G/D 各 1 例, T64N、S70P、K71E 各 2 例, LTNP 组未发现此变异。上述变异两组差异无统计学意义。

B'/C 重组亚型核苷酸序列翻译成氨基酸后与

表1 不同疾病进展情况 HIV/AIDS 患者临床资料

分组	亚型	年龄(岁)	性别	传播途径	CD4 ⁺ T 细胞(个/ μ l)	病毒载量(lg 拷贝/ml)
缓慢进展组	B'(8), B'/C(5)	38 \pm 6	男(9), 女(4)	静脉吸毒(4), 采供血(7), 性传播(2)	624 \pm 180	4.21 \pm 0.75
典型进展组	B'(26), B'/C(9)	40 \pm 9	男(24), 女(11)	静脉吸毒(7), 采供血(19), 性传播(9)	226 \pm 145	5.13 \pm 0.89

注:括号内数据为例数

HXB₂ 毒株、我国较早的流行株 07-97CN001 及 08CN97CNGX_6F 比较(图 2),在缓慢进展组发现 K29L、C37G、R53K、Q54R、H59T、A67G,典型进展组未发现上述变异。在典型进展组发现 A21D、K50R、Q54P、A58T、S62R、K71N,缓慢进展组未发现上述变异。上述变异两组的差异未达到统计学意义。

3. 重要功能区的氨基酸替换:在富半胱氨酸区,本研究发现 B'/C 重组亚型毒株中存在较多的 T23N 变异(7/14),在缓慢进展和典型进展组中分别为 3/5 和 4/9,差异无统计学意义。但该变异只见于云南省来源的病例。在 B'亚型毒株中,来源于缓慢进展和典型进展组的 T23N 各 1 例。T23N(-)与 T23N(+)病例之间病毒载量的差异无统计学意义。6/8 来自云南省的毒株存在 C31S 变异,在缓慢进展和典型进展组中分别为 3/5 和 3/9,差异无统计学意义,该变异同样只见于云南省来源的病例。在缓慢进展和典型进展组中 F31L 变异分别为 3/5 和 4/9,差异无统计学意义。

在核心区,本研究发现 B'亚型毒株中存在较多的 F31L 变异和 I39M、T40K 变异。在缓慢进展和典型进展组差异无统计学意义。核心区的豆蔻酰化位点 GisyGR(44~49)通常极为保守,B'典型进展和缓慢进展组均未发现任何变异。B'/C 重组亚型毒株中只有 1 例 B'/C 典型进展者存在 R50K 变异。

在碱性区保守序列 RKKRRQRRR(49~57),B'

	10	20	30	40	50	60	70
CON S B	MEPVDPRLEP	WKHPGSQPKT	ACTNICYCKKC	CFHCQVCFIT	KGLGISYGRK	KRRQRRRAPQ	DSQTHQVLSL KQ
JC03-SP	-----I. C.
JC12-SP	-----	S. K	S. N.	A. V.
JC14-SP	-----	N.	G.	A.
JC17-SP	-----	S.	A.
LN12-SP	-----	G.	S. N.	A.
JL04-SP	-----	A.
JL12-SP	-----	A.	A.
JL66-SP	-----	R	AN. T.
LN08-TP	-----	L NP	KA. A. I.
LN10-TP	-----	L NP	A.
LN16-TP	-----	N.	L NP	A. T.
LN19-TP N.	TH. A.
LN21-TP	P.	L	H. A.
LN25-TP N.	F.	A.
LN26-TP N.	P	L NK C	I. K	N. A.
LN28-TP N.	G.	S. H.	A.
LN29-TP N.	NP R	H. N. G. P.
LN30-TP N.	G.	A.
LN37-TP G.	A.
LN44-TP	-----	D.
LN48-TP	-----	N. A.
LN50-TP	-----	N. A.
JC11-TP	-----	A	A. A. P.
JC13-TP	-----	K. R.	A.
JC15-TP	-----	R.	T. A.
JC16-TP	-----	H. G. A.
JC20-TP	-----	SN. A.
JC25-TP	-----	I. A.
JC26-TP	-----	HI.	H. A.
JL25-TP	-----	P.	P. A.
JL35-TP	-----	YYS. A.	H. G. A.
JL50-TP	-----	A. A.
JL87-TP	-----	T.
YN04-TP	-----	N.	N. T.

SP 表示缓慢进展组氨基酸序列;TP 表示典型进展组氨基酸序列;国外资料报道的氨基酸所在位置用△标出;只存在于缓慢进展组的氨基酸变异用黑体表示;只存在于典型进展组的氨基酸变异用下划线表示

图1 B'亚型缓慢进展组、典型进展组 Tat 蛋白序列的比较

	10	20	30	40	50	60	70
HXB2	MEPVDPRLEP	WKHPGSQPKT	ACTNICYCKKC	CFHCQVCFIT	KALGISYGRK	KRRQRRRAHQ	NSQTHQASLS KQ
07-97CN001	-----K. S
08CN97CNGX_6F	-----	N.	N. E.	N. R.	SY. L.
LN01-SP	-----	N.
YN15-SP	-----	E. N.	R. SY. L.
YN16-SP	-----	E. N.	R. SY. L.
YN37-SP	-----	E. N.	R. SY. L.
YN104-SP	-----	E. N.	R. SY. L.
LN04-TP K.
LN23-TP N.
LN41-TP	IG. K.
LN43-TP	G. N.
LN45-TP N.
LN51-TP N.
YN08-TP	-----	R.	N. R.	SY. L.
YN102-TP	-----	E. N.	R. SY. L.
YN115-TP	-----	D. D. N.	F. R.	SY. L.

SP 表示缓慢进展组氨基酸序列;TP 表示典型进展组氨基酸序列;国外资料报道的氨基酸所在位置用△标出;只存在于缓慢进展组的氨基酸变异用黑体表示;只存在于典型进展组的氨基酸变异用下划线表示;HXB₂、07-97CN001、08CN97CNGX_6F 为参考株

图2 B'/C 重组亚型缓慢进展组、典型进展组 Tat 蛋白序列的比较

亚型缓慢进展者发生 S53R、S57R 替换各 1 例,B'亚型典型进展者也发生 S53R 1 例。而在 B'/C 重组亚型有 13/14 的毒株中存在 S57R 替换,在缓慢进展和典型进展组中差异无统计学意义。

在富谷氨酰胺区 N6S、Q63E、T64D、A67N、S68L、L69I 在测得序列和 BC 亚型参考株中均常见,反映了亚型特异性。而 A58T 变异在 B'亚型缓慢进展者 1 例,典型进展组 3 例,B'/C 重组亚型毒株中 2 例,在疾病进展情况不同的组间分布无显著性差异,但 A58T(-)病毒载量低于与 A58T(+)病例(P=0.048)。

4. LTNP 组、无症状感染组与 AIDS 组患者 Tat 蛋白二级结构预测结果的比较:以 NNPRIDICT 程序预测我们测得的 Tat 区第一外显子氨基酸序列的二级结构,并进行分组比对(图 3)。在 B'亚型和 B'/C 重组亚型绝大多数患者在 Tat 蛋白的核心区和基础区,尤其是核心区保持了保守的刚性 α -螺旋(5~6 个),但值得注意的是在 B'/C 重组亚型的典型进展组中发现了 5/9 的病例(全部为 CRF07BC)螺旋数仅为 1 个,而 B'/C 重组亚型的缓慢进展组中的 CRF07BC(LN01SP)此区的螺旋数为 5 个。但由于病例数有限,尚未达到统计学差异。在核心区和碱性区交界部位的片层区也相当保守,B'/C 重组亚型片层区要长于 B'亚型,在缓慢进展组合典型进展组都是如此。位于碱性区和富谷氨酰胺区的螺旋结构保守性较低,在缓慢进展组和典型进展组都有缺失的情况发生。

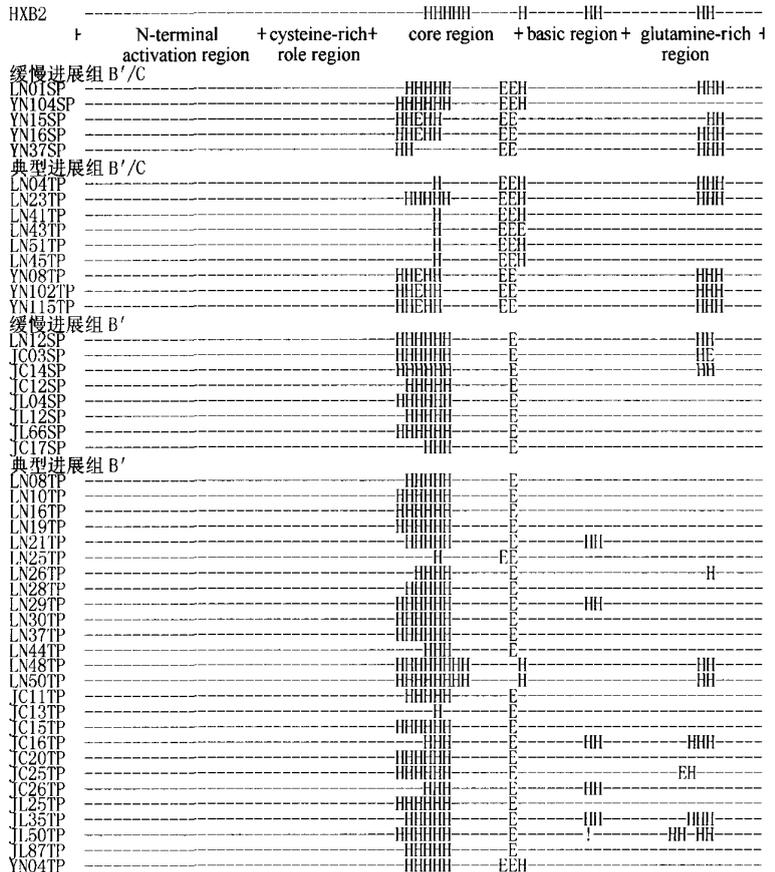
讨 论

HIV-1 的基因组呈现出高度的变异性,这些基因水平的差异当被翻译成有功能的蛋白质后,表现出了生物学性质的差异。根据 Tat 蛋白氨基酸构成和保守区域的分布可分为 6 个区域:N 末端区、富半胱氨酸区、核心区、碱性区、富谷氨酰胺区和 C 末端区。核心区、碱性区和谷氨酰胺丰富区都参与 RNA 的结合。前五区由第一外显子编码,C 末端区则由第二外显子编码。

富半胱氨酸区含有 7 个保守的半胱氨酸残基,该区与反式激活病毒的转录与复制密切相关^[5,6],此区的 T23N 替换在 A、C、D 亚型中常见,该变异能够提高 Tat 对 HIV-1 启动子的激活并使 Tat 结合到细胞激酶正向转录延长因子(positive transcription elongation factor b, P-TEFb),补偿 Tat 蛋白其他部位有害突变,对病毒复制有利,可能促进适应性较差的耐药病毒的复制,利于产生毒力更强病毒^[7]。本研究发

现 B'/C 重组亚型毒株 T23N、C31S、F32Y 变异在缓慢进展和典型进展组中广泛存在,推测与疾病进展情况无关,但该变异只见于云南省来源的 CRF08BC 病例,辽宁省来源的 CRF07 病例则无此变异发生,反映了此变异的亚型特异性。在 B'亚型毒株中,来源于缓慢进展和典型进展组的 T23N 各 1 例。T23N、F32Y 变异此变异并不改变此区半胱氨酸残基的个数,而 C31S 则影响到半胱氨酸残基的个数,其潜在影响还有待进一步研究。

核心区残基构成保守的刚性的螺旋结构,促进 TAT-TAR 结合。Pantano 等研究发现核心区-碱性区的 41~52 位氨基酸高度保守,尤其是 44~50 位氨基酸组成的 turnlike region 保守性更高^[8]。在我国 B'/C 重组亚型毒株中发现在缓慢进展和典型进展组中都存在 S46F 变异,仅见于辽宁省来源的 CRF07BC 毒株,CRF08BC 毒株无此变异,故此变异为亚型特异性变异,其对生物学表型的影响有待进



SP 表示缓慢进展组;TP 表示典型进展组;以 NNPRIDICT 程序预测我们测得的 Tat 区第一外显子氨基酸序列的二级结构未发现显著差异;H 表示螺旋,E 表示折叠,- 表示无规则卷曲

图3 缓慢进展组、典型进展组 Tat 蛋白二级结构预测的比较

一步研究。位于 Tat 28 位和 50 位的赖氨酸残基是乙酰化受体,对病毒复制至关重要。Brès 通过体外实验证明 K28Q 和 K50Q 变异可以引起病毒复制能力丧失, Tat 蛋白对 LTR 的转录激活能力丧失。K28R 会影响病毒的复制能力和对整合的 LTR 的激活, K50R 变异对 Tat 蛋白的转录激活能力无影响,但会使病毒复制延缓 3-4 d^[9]。本研究 1 例 B'/C 重组典型进展者中发现 K50R, 而该患病毒载量处于 10⁵ 的较高水平,并未显示出复制受影响的迹象。

碱性区特异性地与 TAR RNA 的凸起域相互作用,是 Tat 与 TAR 相互作用进行反式激活所必需的。在碱性区保守序列 RKKRRQRRR(49~57)位置, B' 亚型缓慢进展者发生 S53R、S57R 替换各 1 例, B' 亚型典型进展者也发生 S53R 1 例。而在 B'/C 重组亚型有 13/14 的毒株中存在 S57R 替换,显示了此替换的亚型特异性,与疾病进展状况无关。

Tat 区第一外显子特定位置的氨基酸变异可以影响病毒的复制能力,国外学者通过对 HIV-1 B 亚型的体外实验研究发现编码 Tat 蛋白第一外显子的 22、28、41、50 位的基因突变可能使 HIV-1 复制减低,而第 23 位 N 取代 T、58 位 T 取代 A、65 位 D 取代 H 的基因突变,可能使 HIV 复制提高,提示此突变可能与疾病进展有关^[3]。我们在 B' 和 B'/C 缓慢进展组均未发现 22、28、41 位变异,但 1 例 B'/C 重组典型进展者中发现 K50R, 而该患病毒载量处于 10⁵ 的较高水平,并未显出复制受影响的迹象,与体外实验研究结果不同。我们还在云南省的 CRF08BC 毒株中和个别 B' 亚型中发现 T23N, 但未显出与病毒载量以及疾病进展情况的明确关系,而 A58T 虽然在疾病进展情况不同的组间分布无显著性差异,但 A58T(-) 病毒载量低于与 A58T(+) 病例,与国外研究结果一致。

此外我们还发现的一些新的突变: B' 亚型缓慢进展组中发现 K19R、A21S、R57S、A58S、H6N、L69V, B'/C 重组亚型缓慢进展组发现 K29L、C37G、R53K、Q54R、H59T、A67G 的变异,在典型进展组均未发现。B' 亚型典型进展组发现 Q17L、K19N、T20P、A21L/R、A26F/C、I39T、A58P、V67G/D、T64N、S70P、K71E, B'/C 重组亚型典型进展组发现 A21D、K50R、Q54P、A58T、S62R、K71N, 缓慢进展组均未发现,上述变异对 Tat 活性以及病毒复制的影响有待进一步研究。

需要强调的一点是, HIV-1 在体内是呈复杂准种分布的, 由于本实验研究的方法为 PCR 扩增后直接测序, 本研究所发现的 *tat* 基因多态性均代表感染者体内优势准种, 而并不能反映 *tat* 基因多态性的全貌。

本研究试图了解 Tat 蛋白的构象变化对其调节病毒转录能力的影响。我们发现在 B'/C 重组亚型的典型进展组中 5/9 的病例全部为(CRF07BC)此区的螺旋数仅为 1 个, 而 B'/C 重组亚型的缓慢进展组中的 CRF07BC(LN01SP)此区的螺旋数为 5 个。此螺旋数的减少是否与其反式激活活性造成影响还有待进一步实验证实。在核心区 and 碱性区我们也未发现与疾病进展有关的规律性结构改变。交界部位的片层区也相当保守, B'/C 重组亚型片层区要长于 B' 亚型, 但片层区长短与疾病进展情况无关。

上述结果提示我们虽然缓慢进展和典型进展患者 Tat 蛋白的氨基酸序列存在各自的变异特点, 但并未影响其二级结构的改变, HIV 感染后的疾病进展可能与 Tat 蛋白二级结构无明显关系。

参 考 文 献

- Coffin JM. HIV population dynamics in vivo: Implication for genetic variation, pathogenesis, and therapy. *Science*, 1995, 267: 483-489.
- Wei X, Ghosh SK, Talor ME, et al. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature*, 1995, 373: 117-122.
- Pantaleo G, Menzo S, Vaccarezza M, et al. Studies in subjects with long-term nonprogressive human immunodeficiency virus infection. *N Engl J Med*, 1995, 332: 209-216.
- 邢辉, 秦光明, 冯毅, 等. 中国人类免疫缺陷病毒(HIV-1)D 亚型毒株 *gag*, *env* 和 *tat* 基因的序列分析. *中华实验和临床病毒学杂志*, 1999, 13: 157-162.
- Tahtinen M, Strengell M, Collings A, et al. DNA vaccination in mice using HIV-1 *nef*, *rev* and *tat* genes in self-replicating pBN-vector. *Vaccine*, 2001, 19: 2039-2047.
- Boykins RA, Mahieux R, Shankavaram UT, et al. Cutting edge: a short polypeptide domain of HIV-1 Tat protein mediates pathogenesis. *J Immunol*, 1999, 163: 15-20.
- Reza SM, Shen LM, Mukhopadhyay RA. Naturally occurring substitution in human immunodeficiency virus Tat increases expression of the viral genome. *J Virol*, 2003, 77: 8602-8606.
- Pantano S, Carloni P. Comparative analysis of HIV-1 Tat variants. *Proteins*, 2005, 58: 638-643.
- Brès V, Kiernan R, Emiliani S, et al. Tat Acetyl-acceptor Lysines are important for human immunodeficiency virus type-1 replication. *J Biol Chem*, 2002, 277: 22215-22221.

(收稿日期: 2006-05-30)

(本文编辑: 王多春)