

中国南方四省区流行的HIV-1 CRF01_AE病毒株基因特征研究

程春林 冯毅 何翔 林鹏 梁淑家 易志强 贺健梅 胡园园 邢辉
范雁 吴士良 邵一鸣

【摘要】 目的 分析中国南方四省区HIV-1感染者中流行CRF01_AE病毒株的遗传特征。方法 从广东、广西、江西和湖南省(自治区)2006年新报告HIV-1感染者的血浆样本中提取病毒RNA,用反转录/巢式PCR方法扩增gag和env基因片段,对获得的CRF01_AE病毒株核酸序列进行系统进化分析,并通过计算基因距离和Entropy核苷酸多态性差异方法分析毒株的遗传特征。结果 从210例CRF01_AE病毒株感染者中,发现四省区流行的CRF01_AE病毒株存在2个主要的流行簇。流行簇I共有123例样本,未发现与之直接相关的国际参考毒株。流行簇II共有57例样本,与越南CRF01_AE病毒株关系密切,且存在不同时间样本的混杂。gag和env基因遗传距离分析结果表明,流行簇I内基因遗传多样性均明显小于流行簇II;核苷酸多态性分析结果显示,在gag基因片段42个位点核苷酸组成具有显著差异,env基因片段40个位点核苷酸组成存在显著差异;流行簇I相对于流行簇II多态性减少的位点上有61个,多态性增加的位点有21个。结论 在中国南方四省区流行的CRF01_AE病毒株中首次观察到2个独立的流行簇。流行簇I病毒株为该地区最主要的CRF01_AE病毒株,其流行时间相对较短,在人群中所占比例较多,可能是病毒在流行过程中形成的具有传播优势的病毒株。流行簇II病毒与来自于越南的CRF01_AE病毒株有较高同源性,且存在与越南病毒株间的多次传播。

【关键词】 病毒株,HIV-1 CRF01_AE; 流行簇; 进化分析

Genetic characteristics of HIV-1 CRF01_AE strains in four provinces, southern China CHENG Chun-lin*, FENG Yi, HE Xiang, LIN Peng, LIANG Shu-jia, YI Zhi-qiang, HE Jian-mei, HU Yuan-yuan, XING Hui, FAN Yan, WU Shi-liang, SHAO Yi-ming. *National Center for AIDS/STD Control and Prevention, State Key Laboratory for Infectious Disease Prevention and Control, Chinese Center for Disease Control and Prevention, Beijing 100050, China

Corresponding author: SHAO Yi-ming, Email: yshao@bnn.cn

【Abstract】 **Objective** To analyze the genetic characteristics of HIV-1 CRF01_AE strains prevailing in the four provinces of southern China. **Methods** Plasma samples were collected from the newly diagnosed HIV-1 individuals reported in 2006 in Guangdong, Guangxi, Jiangxi and Hunan province. The gag and env gene fragments were amplified from RNA template extracted from plasma using RT and nested PCR methods. CRF01_AE sequences were analyzed by phylogenetic methods and characterized by calculating the genetic distance and Entropy analysis. **Results** Two main epidemic clusters were found to exist in the CRF01_AE strains from 210 HIV-1 CRF01_AE infected individuals collected in the 4 provinces, southern China. It was found that no international reference strain was closely correlated with cluster I, which including 123 samples. The strains in cluster II, consisting 57 cases of samples, were closely related with the strains identified in Vietnam. Genetic distance analysis of gag and env genes showed that the diversity of cluster I was obviously less than that of cluster II. Data on nucleotide polymorphism showed that nucleotides compositions of 42 sites in gag and 40 sites in env were significantly different between the two clusters. When compared with cluster II, the polymorphism decreased at 61 nucleotide sites but increased at 21 sites in cluster I. **Conclusion** This was the first

DOI: 10.3760/cma.j.issn.0254-6450.2009.07.019

基金项目:国家高技术研究发展计划(2006AA02Z418);“艾滋病和病毒性肝炎等重大传染病防治”科技重大专项课题(2008ZX10001-004)

作者单位:100050 北京,中国疾病预防控制中心性病艾滋病预防控制中心传染病预防控制国家重点实验室(程春林、冯毅、何翔、胡园园、邢辉、邵一鸣);广东省疾病预防控制中心(林鹏);广西壮族自治区疾病预防控制中心(梁淑家);江西省疾病预防控制中心(易志强);湖南省疾病预防控制中心(贺健梅);苏州大学医学部生物化学教研室(范雁、吴士良)

程春林、冯毅同为第一作者

通信作者:邵一鸣, Email: yshao@bnn.cn

report describing that two main epidemic clusters were existed in CRF01_AE strains prevailing in the 4 provinces, Southern China. The virus in cluster I was the dominant strain in this region, with shorter period of circulation and higher proportion seen in the HIV-infected population, which might belong to CRF01_AE strain with certain features facilitating the spread of the virus. The virus in cluster II was highly homology with the CRF01_AE strains from Vietnam, and seemed to have had several events of epidemics in populations in border regions of China and Vietnam.

【Key words】 HIV-1 CRF01_AE strains; Epidemic cluster; Phylogenetic analysis

CRF01_AE 病毒株于 1989 年首次从泰国性工作
者中分离^[1],是最早被证实的 HIV-1 流行重组体
(circulating recombinant form, CRF),也是东南亚地
区主要的流行病毒株之一。我国于 1994 年从曾在
泰国从事性工作的一些云南籍妇女中首次检测到该
病毒株^[2]。1996 年 Chen 等^[3]首次发现广西静脉吸
毒和异性性接触人群中存在 CRF01_AE 病毒株流
行。随后广东、江西和湖南等省相继报告检出该病
毒株^[4,5]。目前,CRF01_AE 已成为广西、广东、江西
和湖南四省区 HIV-1 的主要流行株之一^[6]。本研究
对该四省区新报告 HIV-1 感染人群中流行的
CRF01_AE 毒株进行病毒基因序列系统进化分析,
以了解该地区 CRF01_AE 病毒株的遗传变异特征及
其可能来源。

对象与方法

1. 研究对象:在广东、广西、江西和湖南四省区
对 2006 年新报告的 HIV 感染者进行分子流行病学
调查,抽取部分感染者作为研究对象,经知情同意
后,收集血浆样本,病毒亚型确定为 CRF01_AE 的样
本纳入本研究。

2. RNA 模板提取及基因扩增:利用罗氏公司的
MagNa Pure LC 核酸自动提取仪,按照说明书提取
血浆病毒 RNA,通过反转录/巢式 PCR 方法扩增
gag 和 *env* 基因片段。*gag* 基因片段 RT-PCR 扩增引
物为 GAG-L (5'-TCG ACG CAG GAC TCG GCT
TGC-3', 相应于 HXB2 的 686~707) 和 GAG-E2
(5'-TCC AAC AGC CCT TTT TCC TAG G-3', 相应
于 HXB2 的 2032~2011), 反应条件为 50℃ 30 min,
94℃ 3 min; 94℃ 30 s, 55℃ 30 s, 72℃ 90 s, 30 个循
环; 72℃ 10 min; 4℃ 保温。第二轮 PCR 扩增引物
(5'-AGG AGA GAG ATG GGT GCG AGA GCG
TC-3', 相应于 HXB2 的 781~806) 和 GDX (5'-GGC
TAG TTC CTC CTA CTC CCT GAC AT-3', 相应
于 HXB2 的 1861~1836), 反应条件为 94℃ 2 min;
94℃ 30 s, 55℃ 30 s, 72℃ 90 s, 30 个循环; 72℃
10 min; 4℃ 保温。*env* 基因扩增的 RT-PCR 引物为
44F (相应于 HXB2 的 6954~6973) 和 35R (相应于

HXB2 的 7668~7648), 第二轮 PCR 引物为 33F (相
应于 HXB2 的 7002~7021) 和 48R (相应于 HXB2 的
7541~7523), 引物序列与反应条件参见文献[7]。

3. 产物纯化和测序:利用 Axygen 公司的 Axyprep
96 PCR Clean-up Kit 试剂盒,按照说明书对获得的
阳性扩增产物进行纯化。以 GUX 和 GDX 为 *gag* 基
因片段的测序引物,以 33F 和 48R 为 *env* 基因片段的
测序引物。使用 BigDye 3.1 测序试剂盒进行测序反
应,产物纯化后使用 ABI3100 测序仪测定序列。

4. 序列分析:

(1) 应用 Sequencher 4.8 软件对序列进行拼接、
清理和编辑,经 BioEdit 软件进行序列比对后,加入
Los Alamos HIV 序列数据库提供的 HIV-1 分型参考
毒株序列。

(2) 应用 Microsoft Excel 软件中的 CON-
CATENATE 函数将来自于同一个样本的 *gag*、*env* 基
因序列合并成为 *gagenv* 序列。

(3) 应用 MEGA 4.0 软件中的 Maximum
Composite Likelihood 模型计算 *gag* 和 *env* 基因距离
并分别构建 *gag*、*env* 和 *gagenv* 序列的 Neighbor-
Joining 系统进化树,并通过 Interior Branch Test 方法
重复构建进化树 1000 次,用于判断其内部分支的可
靠性^[8];结合获得的 3 个进化树结果判定 CRF01_AE
感染人群中存在的主要流行簇。

(4) 从 Los Alamos HIV 序列数据库下载不同国
家 CRF01_AE 病毒株的 *gag* 和 *env* 参考序列,根据保
留采样时间早和最多有效碱基数的原则挑选序列。
分别与南方四省区 CRF01_AE 病毒株序列构建
Neighbor-Joining 系统进化树,分析不同流行簇的可
能来源。

(5) 应用 Entropy 软件分析不同流行簇之间核苷
酸组成差异^[9]。

结 果

1. 主要流行簇:研究中共获得 210 例
CRF01_AE 感染者作为研究对象,包括 *gag* 基因序
列 194 条, *env* 基因序列 149 条,配对的 *gagenv* 序列
133 条。经比对后进入分析的 *gag* 序列有效长度为

577 bp, *env* 序列有效长度为 234 bp, *gagenv* 序列有效长度为 815 bp。利用 *gag*、*env* 和 *gagenv* 序列构建的系统进化树用于判断研究样本的亚型,其结果如图 1。在不同基因序列构建的 3 个进化树中,从研究对象中获得的序列均与 CRF01_AE 参考病毒株构成稳定的单一进化簇(monophyletic cluster),其 Bootstrap 值为 99%,这证明所研究的 210 份样本均属于 CRF01_AE。其中广东、广西、江西和湖南四省区分别获得 CRF01_AE 样本 75、72、36 和 27 份。在各进化树中未见到明显的地域或危险因素聚集现象。同时,3 个基因片段所构建的系统进化树均形成了 2 个主要的进化簇(进化簇 I 和 II),其中 *gagenv* 序列构建的进化树可靠性更高,进化簇 I 和 II 的可靠性分别为 97% 和 96%。按照 *gag*、*env* 和 *gagenv* 3 个系统进化树中的进化簇分簇结果,可以将样本分成不同的流行簇,其中流行簇 I 样本 123 例(58.6%),流行簇 II 样本 57 例(27.1%)。另外还有其他 30 例样本存在于 2 个主要进化簇以外。

应用 MEGA 软件 Maximum Composite Likelihood 模型计算 2 个主要流行簇 *gag*、*env* 两个基因片段的平均遗传距离(表 1)。结果表明 2 个主要流行簇组间的遗传距离大于任一流行簇组内遗传距离,这一结果进一步证明分簇的合理性。与流行簇 II 相比,流行簇 I 的簇内距离较小。

表 1 CRF01_AE 病毒株 2 个主要流行簇的遗传距离

| 基因片段 | 流行簇 I 组内 | 流行簇 II 组内 | 流行簇 I 和 II 组间 |
|------------|-------------|-------------|---------------|
| <i>gag</i> | 0.013±0.002 | 0.030±0.003 | 0.030±0.005 |
| <i>env</i> | 0.070±0.006 | 0.104±0.009 | 0.135±0.015 |

2. 主要流行簇病毒株来源分析:为了分析四省区 2 个主要流行簇病毒株的可能来源,从 Los Alamos HIV 序列数据库中下载不同国家获得的 CRF01_AE 序列作为参考序列。共选择来自于 25 个国家的 *env* 序列 99 条,与本研究获得的 2 个主要流行簇样本的 *env* 基因序列构建 Neighbor-Joining 系统进化树,结果见图 2。

流行簇 I 是系统进化树中相对独立的进化簇, Bootstrap 值为 98%,该进化簇内无其他国家的序列,与之最接近的参考病毒株为来源于朝鲜的一株病毒(01_AE.KR.94.KR474),其可靠性为 67%;与其进化关系最接近的一组参考序列还包括来自我国云南省、香港地区及新加坡、泰国、美国的参考病毒株。该组序列与流行簇 I 形成一个更大的进化簇,可靠性为 82%。

流行簇 II 的序列与来源于越南不同时间的多条序列形成一个进化簇,可靠性为 91%。不同时间分离的越南病毒株序列与流行簇 II 的序列混杂在一起,形成多个进化分支。

3. 核苷酸多态性分析:应用 Entropy 软件分析 2 个流行簇病毒株的核苷酸多态性差异,结果见图 3(以流行簇 II 序列为参照,浅色线条代表 2 个流行簇核苷酸分布呈显著性差异的位点)。当熵值大于 0 时,表明流行簇 I 在该位点核苷酸组成多态性小于流行簇 II。

gag 基因片段中,有 42 个位点的核苷酸组成存在显著差异(图 3 a)。*env* 基因片段中有 40 个位点的核苷酸组成存在显著差异(图 3 b)。这 82 个位点中,流行簇 I 相对于流行簇 II 多态性减少的位点上有 61 个,多态性增加的位点有 21 个。从总体上看,流行簇 I 的核苷酸组成多态性小于流行簇 II。

4. 不同流行簇病毒株感染样本的危险因素分析:CRF01_AE 不同流行簇病毒株感染样本的最可能感染途径统计结果如表 2。在流行簇 I 中,异性性接触传播占 53.8%,静脉吸毒传播占 35.4%;流行簇 II 中,异性性接触传播占 77.2%,静脉吸毒传播占 14.0%;流行簇 I 和 II 分布人群的组成差异具有统计学意义($P=0.0166$),虽然异性性接触均为 2 个流行簇的主要传播途径,但是静脉注射吸毒感染样本主要分布在流行簇 I 中。

表 2 CRF01_AE 流行病毒株感染途径

| 最可能的传播途径 | 流行簇 I | 流行簇 II | 其他 | 合计 |
|----------|----------|----------|----------|-----------|
| 异性性接触 | 70(56.9) | 44(77.2) | 16(53.3) | 130(61.9) |
| 静脉注射吸毒 | 46(37.4) | 8(14.0) | 7(23.3) | 61(29.0) |
| 男男性行为 | 0(0.0) | 0(0.0) | 3(10.0) | 3(1.4) |
| 输血/血制品 | 1(0.8) | 1(1.8) | 1(3.3) | 3(1.4) |
| 传播途径不详 | 6(4.9) | 4(7.0) | 3(10.0) | 13(6.2) |

注:括号外数据为例数,括号内数据为百分率(%)

讨 论

利用系统进化分析方法确定的进化簇在很大程度上代表了样本间的遗传学和流行病学关联,具有重要的流行病学意义。在确定进化簇时,除了需要选择合适的模型构建系统进化树以外,还需根据分析目的选择合适的检验方法。在 HIV 分子流行病学研究中,常将系统进化树上聚集成簇的且检验值 > 70% 的分支定义为可靠进化簇。

为了在系统进化分析中获得更为可靠的结果,本研究利用 Interior Branch Test 方法检验

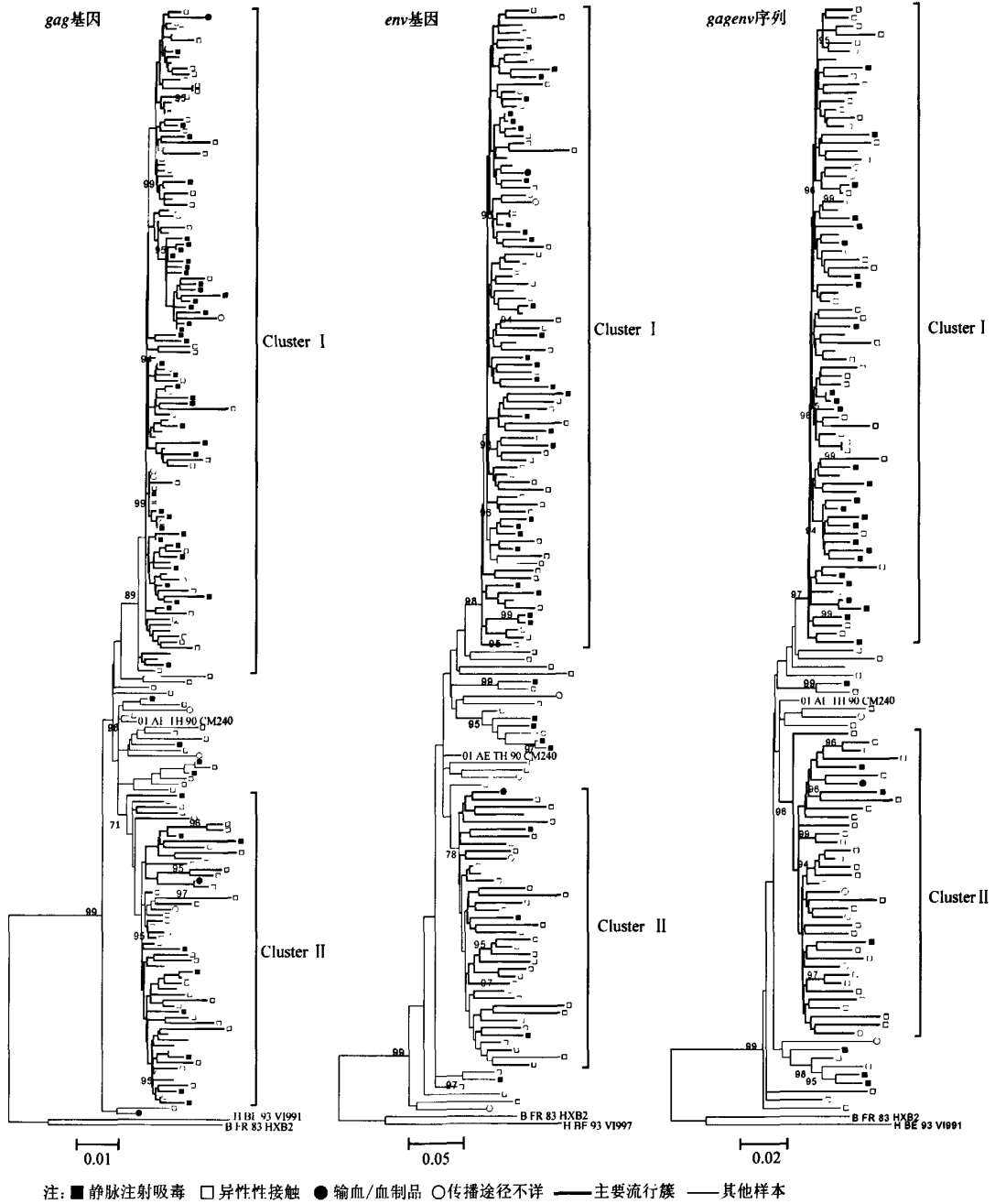
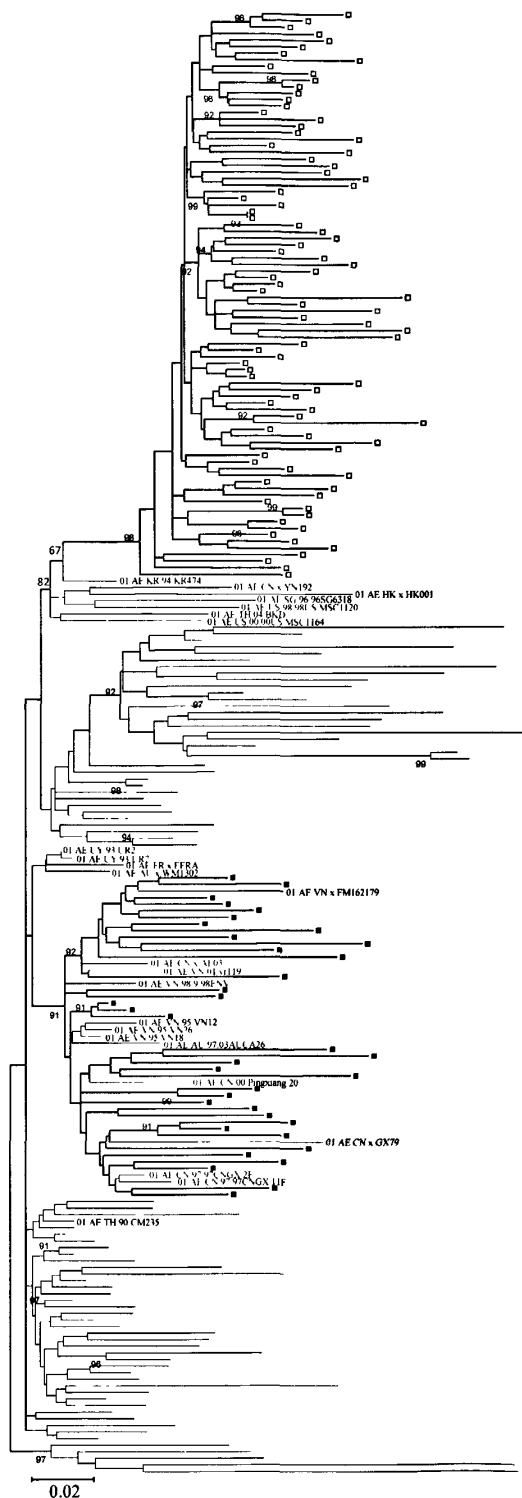


图1 CRF01_AE病毒株不同基因序列的系统进化分析

Neighbor-Joining系统进化树分支的可靠性,提高了系统进化方法在检验同一病毒亚型病毒株在亚种以下水平的分辨率。同时,将 *gag*、*env* 序列拼接成 *gagenv* 序列后进行系统进化分析,可以弥补因使用较短序列、遗传信息不够造成的系统进化树不够稳定的缺陷,使结果更加稳定可靠。综合以上两种分析方法,在我国南方四省区观察到CRF01_AE病毒株存在2个主要流行簇,簇内感染者占所有研究对

象的85.7%,同时,仍有部分样本未包含于这2个流行簇之中,表明该地区的CRF01_AE病毒株可能有3个或以上的病毒株来源。

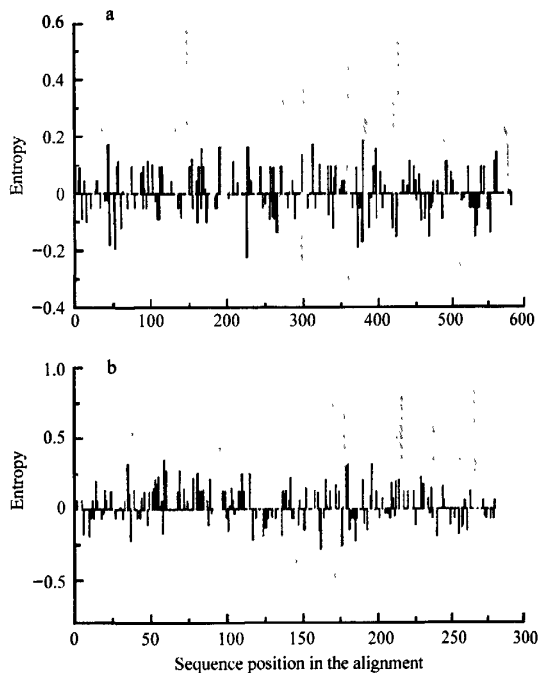
对2个主要流行簇病毒株的来源分析表明,流行簇II的 *env* 序列与不同时间采集的越南病毒株序列混杂在一起(图2),也有部分研究显示我国广西地区流行的CRF01_AE病毒株与越南病毒株存在较高的遗传同源性^[10-13],这些结果表明,由于我国与越



注：— 我国南方四省区序列；— 参考序列；□ 进化簇 I；■ 进化簇 II

图2 主要流行簇病毒株的来源分析

南边境地区人员交流频繁,造成流行簇 II 内的病毒株可能在两国边境人群中存在多次相互传播过程。



注：a: gag 基因片段 Entropy 分析结果；b: env 基因片段 Entropy 分析结果；线段长度表示该点核苷酸多态性差异大小,浅色线条表示核苷酸多态性有显著性差异的位点

图3 我国南方四省区 CRF01_AE 病毒株 2 个主要流行簇核苷酸多态性差异分析

对流行簇 I 的进化分析结果表明, gag 和 env 基因片段簇内遗传距离的结果支持簇 I 流行时间较簇 II 晚(表1); Entropy 核酸多态性分析结果也显示流行簇 I 多数核苷酸位点的多态性小于流行簇 II (图3), 这些结果表明流行簇 I 相对流行时间较短。虽然簇 I 发生流行的时间较晚,但在本研究中,流行簇 I 占所有研究对象的比例明显高于流行簇 II 和其他病毒株(表2),表明该病毒株在传播过程中可能存在一定优势,进而在较短时间内成为四省区 CRF01_AE 的优势病毒株。

对于2个主要流行簇危险因素的分析结果表明,静脉吸毒传播在流行簇 I 中占有更重要的地位,表明流行簇 I 在人群中的优势可能是由于快速传播引起。在不同基因或包括不同参考序列的系统进化树中,流行簇 I 均独立而稳定的存在,未发现与之明确相关的已知病毒株。因此,推测流行簇 I 病毒株是由于病毒进化或快速传播而形成的新流行病毒株,有可能该病毒株在进化过程中获得了某些生物学方面的优势,使之在相对短的时间内造成快速流行,但影响该病毒株传播优势的生物学因素还有待于进一步研究证实。

参 考 文 献

- [1] McCutchan FE, Hegerich PA, Brennan TP, et al. Genetic variants of HIV-1 in Thailand. *AIDS Res Hum Retroviruses*, 1992, 8:1887-1895.
- [2] Cheng H, Zhang J, Capizzi J, et al. HIV-1 subtype E in Yunnan, China. *Lancet*, 1994, 344:953-954.
- [3] Chen J, Young NL, Subbarao S, et al. HIV type 1 subtypes in Guangxi province, China, 1996. *AIDS Res Hum Retroviruses*, 1999, 15:81-84.
- [4] 冯铁建, 邵一鸣, 李良成, 等. 深圳地区 HIV-1 流行毒株的 *env* 基因序列测定和亚型分析. *中华传染病学杂志*, 1998, 16:203.
- [5] 邢辉, 潘品良, 苏玲, 等. 1996—1998 年中国流行的 E 亚型艾滋病病毒 1 型毒株的分子流行病学研究. *中国性病艾滋病防治*, 2002, 8:201-203.
- [6] 邢辉, 梁浩, 万卓越, 等. 中国 CRF01_AE 亚型人类免疫缺陷病毒毒株的分子流行病学研究. *中华预防医学杂志*, 2004, 38:300-304.
- [7] Leung TW, Mak D, Wong KH, et al. Molecular epidemiology demonstrated three emerging clusters of human immunodeficiency virus type 1 subtype B infection in Hong Kong. *AIDS Res Hum Retroviruses*, 2008, 24:903-910.
- [8] Tamura K, Dudley J, Nei M, et al. Molecular evolutionary genetic analysis (MEGA) software version 4.0. *Mol Biol Evol*, 2007, 24:1596-1599.
- [9] <http://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy.html>.
- [10] Yu XF, Chen J, Shao Y, et al. Emerging HIV infections with distinct subtypes of HIV-1 infection among injection drug users from geographically separate locations in Guangxi province, China. *J Acquir Immune Defic Syndr*, 1999, 22:180-188.
- [11] Quan VM, Chung A, Long HT, et al. HIV in Vietnam: the evolving epidemic and the prevention response, 1996 through 1999. *J Acquir Immune Defic Syndr*, 2000, 25:360-369.
- [12] Kato K, Shiino T, Kusagawa S, et al. Genetic similarity of HIV type 1 subtype E in a recent outbreak among injecting drug users in northern Vietnam to strains in Guangxi province of southern China. *AIDS Res Hum Retroviruses*, 1999, 15:1157-1168.
- [13] Kato K, Kusagawa S, Motomura K, et al. Closely related HIV-1 CRF01_AE variant among injecting drug users in northern Vietnam: evidence of HIV spread across the Vietnam-China border. *AIDS Res Hum Retroviruses*, 2001, 17:113-123.

(收稿日期:2009-02-26)

(本文编辑:张林东)

· 疾病控制 ·

一起肺炎支原体和腺病毒混合感染暴发疫情的调查

王全楚 段新科 李争

2009 年 1 月 4 日至 2 月 14 日, 驻豫某部队连续出现以发热、头痛、咽痛、咳嗽和肺部感染等临床症状的病例。经流行病学调查和临床、实验室检查证实, 是由腺病毒和肺炎支原体混合感染引起的暴发流行。此期间共发生 122 名病例, 罹患率 24.6%。发病时间集中在 1 月 24 日和 2 月 12 日两个发病高峰, 呈双峰型, 共发病 102 例, 占发病人数的 83.6%。病例全部集中在一个院落, 其他单位均无病例发生。患者既有老兵又有新兵, 发病年龄 18~25 岁, 无明显年龄差异。122 例中临床表现为发热 10 例 (8.2%)、头痛、头晕 114 例 (93.4%)、咽痛 32 例 (26.2%)、胸闷 24 例 (19.6%)、肌肉酸痛、关节疼痛、腰痛 37 例 (30.3%), 继发结膜炎 (36.8%) 和腹痛、腹泻 (33.6%)。13 例重症肺炎患者 X 线均出现不同程度肺实变, 3 例伴胸腔积液。检测 122 例患者血清腺病毒 IgM 抗体, 阳性率 70%。取 13 例重症肺炎患者血清和咽拭子标本送解放军第三〇二医院病毒实验室检测。荧光抗体染色有 6 例

发现腺病毒。7 例血清支原体 MP IgM 抗体阳性, 结合临床特征和微生物培养鉴定, 确定此次疫情是由腺病毒和肺炎支原体混合感染所致。

讨论: 呼吸道感染的病原体较复杂, 有可能由两种完全不同病原体引起或混合感染并发。本组发病人员分布在同一个营院, 接触频繁, 通风不畅, 为疫情蔓延提供了条件。首批患者出现后, 未采取积极的隔离、消毒措施, 随后病例大量出现, 是引起暴发的主要原因。此外, 新兵训练强度大, 训练时间长, 对气候变化的抵抗能力下降, 也是造成集体暴发的主要诱因。感染腺病毒肺炎后的临床表现与支原体肺炎极为相似, 一般病情较轻, 重者可有大量胸腔积液。此病主要见于新兵, 似与兵营的训练疲惫和居住拥挤的特殊环境有关。支原体肺炎是一种自限性疾病, 使用抗生素治疗后虽可缩短病程, 减少并发症的发生, 但消除微生物效果并不理想。本次通过分子生物学及血清学多种方法证实了在本次疫情中有两种病原体同时存在, 为处理疫情提供了完整准确的实验室依据。

(收稿日期:2009-03-01)

(本文编辑:张林东)

DOI: 10.3760/cma.j.issn.0254-6450.2009.07.020

作者单位: 450042 郑州, 解放军第一五三医院(王全楚); 解放军第一五二医院(段新科、李争)