

## · HIV 感染及基因分型 · 论著 ·

# 中国 HIV CRF01\_AE 亚型及其相关重组亚型包膜糖蛋白分子进化及共识序列研究(2015—2021 年)

胡兆会<sup>1</sup> 张娟<sup>2</sup> 陈凤<sup>3</sup> 索家乐<sup>1</sup>

<sup>1</sup>成都市武侯区疾病预防控制中心检验科, 成都 610093; <sup>2</sup>云南省疾病预防控制中心消毒病媒所, 昆明 650034; <sup>3</sup>大理白族自治州血吸虫病防治研究所, 大理 671000

通信作者: 索家乐, Email: suojiiale1997@foxmail.com

**【摘要】** 目的 了解我国 HIV CRF01\_AE 亚型及其相关重组亚型的包膜糖蛋白 (envelope glycoprotein, ENV) 分子流行情况, 建立 ENV 共识序列以了解氨基酸突变及蛋白质三级结构。方法 收集 Los Alamos 数据库中 2015—2021 年我国 HIV-1 CRF01\_AE 亚型及其相关重组亚型的完整 ENV 基因序列, 构建系统发育进化树并建立 ENV 共识序列, 使用 ProtScale 及 PROVEAN 对 ENV 共识序列广谱中和抗体表位中的氨基酸亲疏水性变化及有害性突变进行预测, 同时分析氨基酸突变对 ENV 三级结构的影响。结果 系统发育分析结果显示, 有 70.0% (14/20) 的 CRF01\_AE 亚型序列属于进化簇 4, 在 ENV 共识序列的广谱中和抗体表位 V1/V2 及 Loop D 中分别存在 4 个和 1 个的氨基酸突变使亲疏水性发生变化, PROVEAN 预测结果显示 V1/V2 位点的 G152D 及  $\beta$ 23/V5/ $\beta$ 24 表位的 M476I 为有害突变。ENV 共识序列的蛋白质三级结构 pLDDT 值为 86.2, 在 V1/V2 表位存在由于有害突变所产生的新  $\alpha$  螺旋。结论 我国 HIV-1 CRF01\_AE 亚型基因谱系众多, 加剧了重组病毒的形成。在 ENV 共识序列 V1/V2 结构域中由于氨基酸突变使其蛋白空间结构发生改变, 这种改变可能会对后续广谱中和抗体的诱导产生影响。

**【关键词】** HIV; CRF01\_AE; 共识序列; 广谱中和抗体; 分子进化; 系统发育分析

基金项目: 云南省高层次卫生健康技术人才培养专项经费(H-2018072)

DOI: 10.3760/cma.j.cn331340-20240228-00041

## Molecular evolution and consensus sequence research of HIV CRF01\_AE subtype and its related recombinant subtype envelope glycoprotein in China (2015-2021)

Hu Zhaohui<sup>1</sup>, Zhang Juan<sup>2</sup>, Chen Feng<sup>3</sup>, Suo Jiiale<sup>1</sup>

<sup>1</sup>Clinical Laboratory, Wuhou District Center for Disease Control and Prevention, Chengdu 610093, China; <sup>2</sup>Institute of Disinfection and Vector Control, Yunnan Provincial Center for Disease Control and Prevention, Kunming 650034, China; <sup>3</sup>Dali Autonomous Prefecture Institute of Schistosomiasis Control and Prevention, Dali 671000, China

Corresponding author: Suo Jiiale, Email: suojiiale1997@foxmail.com

**【Abstract】 Objective** To understand the molecular epidemiology of HIV CRF01\_AE subtype and its related recombinant subtype envelope glycoprotein (ENV) in China, and to establish ENV consensus sequences to analyze amino acid mutations and protein tertiary structure. **Methods** Full-length ENV gene sequences of HIV-1 CRF01\_AE subtype and related recombinant subtypes in China from 2015 to 2021 were collected from the Los Alamos database. A phylogenetic evolutionary tree was constructed, and ENV consensus sequences were established. ProtScale and PROVEAN were used to predict changes in amino acid hydrophobicity in broad neutralizing antibody epitopes of the ENV consensus sequence and harmful mutations, and the impact of amino acid mutations on the tertiary structure of the ENV protein was analyzed. **Results** Phylogenetic analysis showed that 70.0% (14/20) of CRF01\_AE subtype sequences belonged to evolutionary cluster 4. In the broad neutralizing antibody epitope V1/V2 and Loop D of the ENV consensus sequence, there were 4 and 1 amino acid mutations, respectively, causing changes in hydrophobicity. PROVEAN predicted that the mutations G152D at the V1/V2 and M476I at the  $\beta$ 23/V5/ $\beta$ 24 epitopes were harmful.

The protein tertiary structure pLDDT value of the ENV consensus sequence was 86.2, with a new  $\alpha$  helices generated in the V1/V2 epitope due to harmful mutations. **Conclusions** The genetic lineage of HIV-1 CRF01\_AE subtype in China is diverse, contributing to the formation of recombinant viruses. Amino acid mutations in the V1/V2 structural domain of the ENV consensus sequence result in changes in its protein spatial structure, which may impact the subsequent induction of broad neutralizing antibodies.

**【Key words】** HIV; CRF01\_AE; Consensus sequence; Broadly neutralizing antibody; Molecular evolution; Phylogenetic analysis

**Fund program:** Special Fund for Training High-Level Health Technical Personnel in Yunnan Province (H-2018072)

DOI: 10.3760/cma.j.cn331340-20240228-00041

截至 2020 年, 中国有 105 万人感染 HIV<sup>[1-3]</sup>。经抗逆转录病毒治疗使 AIDS 成为一种可控制的慢性病, 疫苗的研发或许能有效控制 HIV 传播<sup>[4]</sup>。CRF01\_AE 亚型自 20 世纪 90 年代初传入我国后快速传播, 目前已成为我国流行最广的亚型, 且在传播过程中 CRF01\_AE 又作为母株参与形成了其他 HIV-1 重组毒株<sup>[5-7]</sup>。HIV 感染人体主要通过包膜糖蛋白(envelope glycoprotein, ENV)的 GP120 和 GP41 与细胞表面的 CD4 和共受体 CCR5 或 CXCR4 相互作用完成的, 同时 ENV 也是病毒表面介导病毒与细胞膜融合的唯一靶点<sup>[8-9]</sup>。对 ENV 进行设计, 选择与当前流行株最为接近的共识序列 (consensus sequence, CONS), 使其诱导对异源 HIV-1 毒株具有保护作用的广谱中和抗体是 HIV-1 疫苗开发的主要目标, 其中针对 V2 可变环的高水平抗体是控制病毒传播的关键<sup>[4,10]</sup>。本研究以 2015—2021 年中国地区 HIV CRF01\_AE 亚型及其相关重组亚型的 ENV 序列为样本, 进行系统发育分析并构建共识序列, 同时对共识序列广谱中和抗体位点的氨基酸序列及空间结构进行分析, 为我国后续的 HIV 疫苗免疫原的设计与改造提供参考。

## 材料与方法

### 一、材料来源

收集 2015—2021 年上传至 Los Alamos 数据库中我国 HIV CRF01\_AE 亚型及其相关重组亚型的完整 ENV(HXB2 6225bp-8795bp)序列, 以 HIV 样本收集的时间及省份对序列进行命名, 共 84 条序列,

其中 CRF01\_AE 亚型 20 条 (20/84, 23.8%)。引入 CRF01\_AE 的 7 个不同发育簇的各 3 条序列作为参考<sup>[11]</sup>, HXB2/K03455 为 2002 年发布的 HIV 参考序列。

### 二、研究方法

用 Mega7.0 软件使用 N-J 法构建进化树, 并用 itol 在线软件对进化树进行美化。使用在线工具 (<https://www.hiv.lanl.gov/content/sequence/CONSENSUS/Consensus.html>) 构建我国 CRF\_01AE 及其相关重组亚型的 ENV 共识序列, 并使用 Bioedit 7.0 对共识序列与参考株 HXB2 进行氨基酸序列一致性比对。对广谱中和抗体表位 Inner domain、V1/V2、Loop D、 $\beta$ 23/V5/ $\beta$ 24、MPER 内的氨基酸突变进行分析, 使用 ProtScale 在线工具对共识序列氨基酸突变所引起的亲疏水性变化进行分析。使用 PROVEAN 工具评估突变对于 ENV 结构及功能的影响, 在 PROVEAN 中, 使用了 -2.5 的阈值 ( $\leq -2.5$  认为有害,  $> -2.5$  认为中性)<sup>[12]</sup>。同时使用 AlphaFold 在线工具对 ENV 共识序列及参考序列 HXB2 ENV 的 V1V2 结构域进行 3D 建模, 计算 pLDDT 值 (衡量单个残基置信度指标), 当 pLDDT 值  $> 70$  时则该模型是一个普遍正确的骨架预测<sup>[13]</sup>, 并使用 Pymol 软件对其进行比对。

## 结 果

### 一、进化树构建

进化树结果见图 1, 其中 CRF01\_AE 亚型中 70.0% (14/20) 的序列位于进化簇 4 中, 2018Jiangxi/MT611503、2015Hubei/MG655206 位于进化簇 1 中,

2016Yunnan/MT624757 位于进化簇 6 中 (自举值大于 90)。其余 3 条 CRF01\_AE 亚型序列(2015Hubei/MG655211、2015Hubei/MG655218、2016Yunnan/MT624756)在进化树上与进化簇 1、5 靠近但均未形成独立进化分支 (自举值<90)。此外 CRF114\_0155 的

ENV 片段在进化发育树中与进化簇 4 聚集在一起, CRF112\_01B 及 CRF104\_0107 的 ENV 片段与进化簇 5 聚集在一起, 提示这些重组亚型可能是由 CRF01\_AE 亚型的进化簇 4、5 参与重组形成, 进化簇 4、5 主要在我国北方的男男性接触人群间传播<sup>[1]</sup>。同时

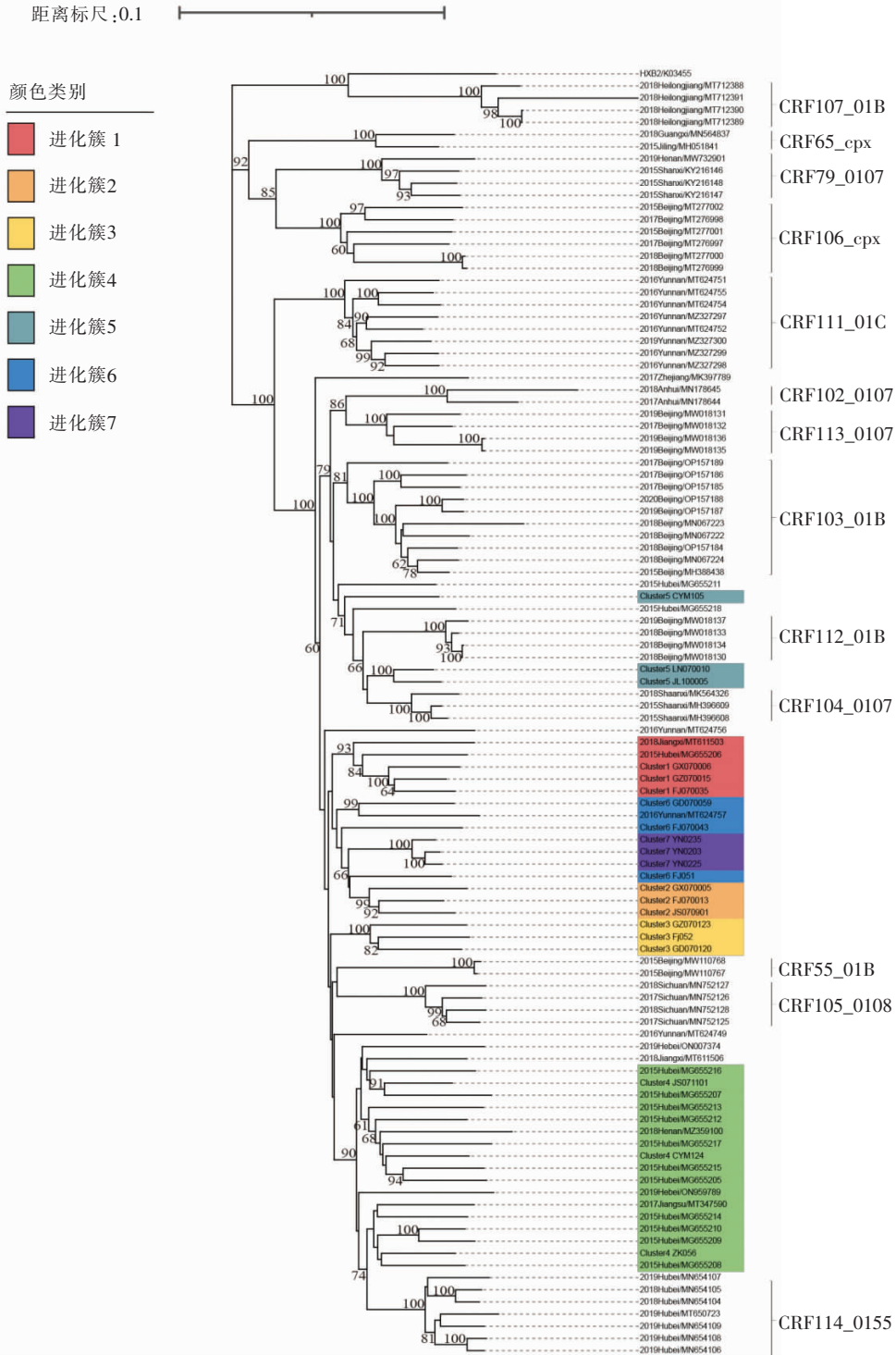


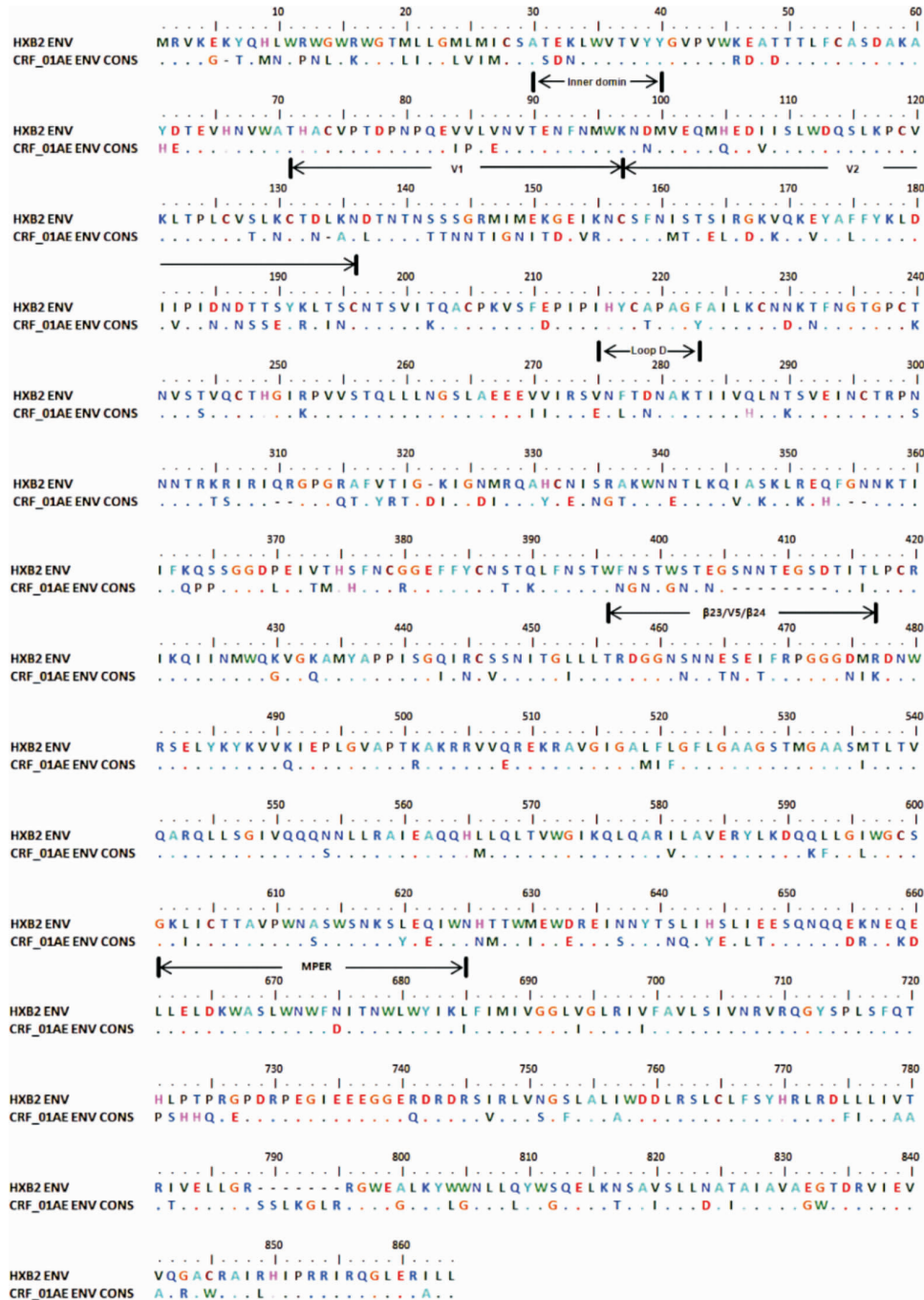
图 1 基于 CRF01\_AE 亚型及其相关重组亚型的完整包膜糖蛋白基因序列构建的 N-J 法系统发育进化树

亚型 CRF65\_cpx,CRF79\_0107,CRF106\_cpx 的 ENV 片段在进化树中互相靠近聚集在同一进化分支中。

### 二、氨基酸突变分析

通过将 ENV 共识序列与参考序列 HXB2 进行氨基酸比对发现,在 ENV 共识序列广谱中和抗体表位 Inner domin (AA90~100)、V1/V2 (AA131~191)、

Loop D(AA275~283)、 $\beta$ 23/V5/ $\beta$ 24(AA451~477)、MPER(AA661-685)中分别存在 1、33、3、7、2 个氨基酸突变,如图 2 所示,其中 V1/V2 表位及 Loop D 表位中分别有 4 个(E150I、S162T、I165L、D185N)和 1 个(V275E)氨基酸突变导致了蛋白水性发生改变。经 PROVEAN 预测发现,ENV 共识序列存在 2 个有害



注:图中分别标注出 Inner domain、V1/V2、Loop D、 $\beta$ 23/V5/ $\beta$ 24、MPER 抗原表位;(·)表示氨基酸序列一致;(-)表示由于优化对其引入新的间隙,氨基酸位置编号均以 HXB2 株为参考

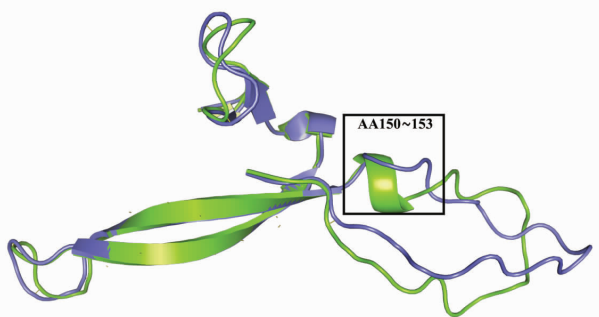
图 2 建立 CRF01\_AE 亚型及其相关重组亚型的包膜糖蛋白共识序列并与参考株 HXB2 进行氨基酸序列比对



的氨基酸突变,分别为 V1/V2 表位的 G152D 以及  $\beta$ 23/V5/ $\beta$ 24 表位的 M475I。

### 三、蛋白结构分析

使用 Alphafold 对 ENV 共识序列与参考株 HXB2 的 V1/V2 结构域的蛋白质三级结构进行建模,pLDDT 值分别为 86.2 与 83.3,模型可信。经 Pymol 进行蛋白空间结构比对发现,在 AA150~153 处 ENV 共识序列存在由于氨基酸突变所产生的新  $\alpha$  螺旋,如图 3 所示,RMSD 值为 0.590,RMSD 值越大即蛋白质之间结构差异越大。



注:绿色为包膜糖蛋白共识序列的 V1/V2 表位蛋白结构,蓝紫色为 HXB2 的 V1/V2 表位蛋白结构,AA150~153 为氨基酸 150 至 153,氨基酸位置编号均以 HXB2 株为参考

图 3 包膜糖蛋白共识序列与参考株 HXB2 V1/V2 表位蛋白结构比对图

## 讨 论

HIV CRF01\_AE 亚型起初在中国吸毒者和性活跃人群中传播,随着近年来网络社会的迅速发展,导致人们更容易通过社交媒体及网络应用相互联系,其流行也由高危人群向一般人群扩散<sup>[14-15]</sup>。

### 一、我国 CRF01\_AE 亚型有多个重组亚型

本研究通过对数据库中 2015—2021 年 CRF01\_AE 亚型及其相关重组亚型的 ENV 序列系统发育分析发现,有 3 条 CRF01\_AE 亚型的序列不属于任何一支进化簇,但却在进化树中与进化簇 1、进化簇 5 靠近,这可能是由于在 CRF01\_AE 的亚型内重组所导致,在我国的 CRF01\_AE 流行疫情中,有高达 54.9% 的序列显示为亚型内重组体<sup>[16]</sup>。同时有 3 个循环重组亚型(CRF65\_cpx、CRF79\_0107、CRF106\_cpx)在进化树中相互靠近并处于同一进化分支,猜测这 3 种亚型均由 CRF01\_AE 亚型同一进

化簇发生重组形成。我国 CRF01\_AE 亚型由多个不同基因谱系组成,这些谱系的广泛流行加剧了新的重组病毒形成,随后这些重组病毒在不同的危险群体和地理区域中传播。因此对中国 CRF01\_AE 谱系及其相关重组亚型的研究和监测将对制定有效的干预计划提供帮助<sup>[11]</sup>。

二、在 HIV 的 V1/V2 表位及 Loop D 表位中存在氨基酸突变所导致的蛋白水性变化

病毒重组是 HIV 重要的进化策略,同时也是疫苗设计的阻碍,不同亚型人群在合并感染期间的二次接触,将导致全世界重组体的增加<sup>[17]</sup>。病毒的重组最终可能会导致两种结果,一是为病毒新的适应提供更多潜力,另一种是使 AIDS 人群的病毒基因变得同质化<sup>[18]</sup>。本研究以 CRF01\_AE 亚型及其相关重组亚型的完整 ENV 序列为样本构建共识序列,将 ENV 共识序列与当前流行株之间的距离降至最低<sup>[19]</sup>。通过对所建立的 ENV 共识序列与参考株 HXB2 氨基酸序列进行比对及预测发现:在 V1/V2 表位及 Loop D 表位中存在有氨基酸突变所导致蛋白水性变化,V1/V2 是广谱中和抗体 PG9 识别的主要糖肽表位,但也是 HIV 逃避免疫的主要区域,HIV 通过 V1 结构域的氨基酸变异及糖基化来逃避中和抗体的识别<sup>[20]</sup>。同时在 Loop D 表位上发生的部分氨基酸突变会对 VRC01 类抗体产生较大的空间位阻,从而使毒株产生耐药性<sup>[21]</sup>,分析这些广谱中和抗体表位中的氨基酸对于 HIV 疫苗设计至关重要。另外 PROVEAN 预测结果显示存在 2 个有害突变 G152D、M476I 分别位于 V1/V2 表位以及 CD4 结合位点  $\beta$ 23/Loop V5/ $\beta$ 24 中,这些有害性突变可能会使 HIV 病毒毒力减弱,这是由于病毒长期流行适应人体环境,为获得免疫逃逸的结果<sup>[7]</sup>。

三、V1/V2 结构域的复杂性影响广谱中和抗体的诱导

由 3 个 ENV 构成的类天然三聚体不仅可以更真实地模拟病毒与免疫系统相互作用的过程,同时也能呈现多个广谱中和抗体表位,并阻碍非中和抗体表位的暴露<sup>[22]</sup>,V1/V2 作为 ENV 三聚体结合结构域,三聚体构象受到 V1 和 V2 的序列、关键残基的

取代、原子及原子间的作用力影响。在 ENV 共识序列 AA150~153 发现的蛋白结构变化也与氨基酸突变位点 E150I 及 G152D 的预测结果相印证。新的  $\alpha$  螺旋可能会对 ENV 三聚体构象产生影响,因此了解 V1/V2 结构域的复杂性能进一步了解广谱中和抗体发生的诱导变化<sup>[23]</sup>,为 HIV 疫苗研究提供参考。

**利益冲突** 所有作者均声明不存在利益冲突

**作者贡献声明** 胡兆会:数据收集、分析及文章撰写;张娟:数据收集、分析及文章撰写;陈凤:研究指导、工作支持;索家乐:序列分析、文章撰写

### 参 考 文 献

- [1] 何纳. 中国艾滋病流行病学研究新进展[J]. 中华疾病控制杂志, 2021, 25(12): 1365-1368, 1480. DOI: 10.16462/j.cnki.zhjbkz.2021.12.001.
- [2] Wang T, Gu Y, Ran L, et al. Ways of HIV transmission in China: The effect of age, period, and cohort[J]. *Front Public Health*, 2022, 10: 941941. DOI: 10.3389/fpubh.2022.941941.
- [3] Laurence J. "The Path That Ends AIDS": National and international politics in conflict with treatment and prevention science[J]. *AIDS Patient Care STDS*, 2023, 37(11): 505-506. DOI: 10.1089/apc.2023.29020.com.
- [4] 王明明. HIV-1 进化阶段性包膜蛋白三聚体序贯免疫诱导广谱中和抗体的研究[D]. 长春: 吉林大学, 2023.
- [5] 王晓林, 贾磊, 李韩平, 等. 1996-2014 年中国 HIV-1 毒株 CRF01\_AE 亚型传播簇和传播网络研究[J]. *中华流行病学杂志*, 2019, 40(1): 84-88. DOI:10.3760/cma.j.issn.0254-6450.2019.01.017.
- [6] 袁子焰, 杨渊, 殷环, 等. 亚洲 HIV-1 基因型的流行特征(1990—2020 年)[J]. *国际流行病学传染病学杂志*, 2022, 49 (2): 91-96. DOI: 10.3760/cma.j.cn331340-20211206-00234.
- [7] 李敬云. 艾滋病病毒 1 型 CRF01\_AE 毒株的过去、现在和将来[J]. *中华流行病学杂志*, 2016, 37(4): 443-449. DOI: 10.3760/cma.j.issn.0254-6450.2016.04.001.
- [8] Xiao T, Cai Y, Chen B. HIV-1 entry and membrane fusion inhibitors[J]. *Viruses*, 2021, 13(5): 735. DOI: 10.3390/v13050735.
- [9] 沈鸿霖, 张芝晴, 李少伟, 等. HIV-1 Env 三聚体抗原改造研究进展[J]. *生物工程学报*, 2020, 36(1): 25-32. DOI: 10.13345/j.cjb.190162.
- [10] Haynes BF, Wiehe K, Borrow P, et al. Strategies for HIV-1 vaccines that induce broadly neutralizing antibodies[J]. *Nat Rev Immunol*, 2023, 23 (3): 142-158. DOI: 10.1038/s41577-022-00753-w.
- [11] Feng Y, He X, Hsi JH, et al. The rapidly expanding CRF01\_AE epidemic in China is driven by multiple lineages of HIV-1 viruses introduced in the 1990s[J]. *AIDS*, 2013, 27(11): 1793-1802. DOI: 10.1097/QAD.0b013e328360db2d.
- [12] Choi Y, Chan AP. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels[J]. *Bioinformatics*, 2015, 31(16): 2745-2747. DOI: 10.1093/bioinformatics/btv195.
- [13] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold[J]. *Nature*, 2021, 596(7873): 583-589. DOI: 10.1038/s41586-021-03819-2.
- [14] Pagkas-Bather J, Young LE, Chen YT, et al. Social network interventions for HIV transmission elimination[J]. *Curr HIV/AIDS Rep*, 2020, 17(5): 450-457. DOI: 10.1007/s11904-020-00524-z.
- [15] Zhang J, Fan Q, Luo M, et al. Phylogenetic evidence of HIV-1 transmission linkage between two men who have sex with men[J]. *Virology*, 2021, 18(1): 106. DOI: 10.1186/s12985-021-01573-5.
- [16] Jia L, Gui T, Li L, et al. A considerable proportion of CRF01\_AE strains in China originated from circulating intrasubtype recombinant forms (CIRF)[J]. *BMC Infect Dis*, 2015, 15:528. DOI: 10.1186/s12879-015-1273-5.
- [17] Vuilleumier S, Bonhoeffer S. Contribution of recombination to the evolutionary history of HIV[J]. *Curr Opin HIV AIDS*, 2015, 10(2): 84-89. DOI: 10.1097/COH.000000000000137.
- [18] Song H, Giorgi EE, Ganusov VV, et al. Tracking HIV-1 recombination to resolve its contribution to HIV-1 evolution in natural infection [J]. *Nat Commun*, 2018, 9 (1): 1928. DOI: 10.1038/s41467-018-04217-5.
- [19] Meyerhoff RR, Scarce RM, Ogburn DF, et al. HIV-1 consensus envelope-induced broadly binding antibodies[J]. *AIDS Res Hum Retroviruses*, 2017, 33(8): 859-868. DOI: 10.1089/AID.2016.0294.
- [20] Shen X, Laher F, Moodie Z, et al. HIV-1 vaccine sequences impact V1V2 antibody responses: A comparison of two poxvirus prime gp120 boost vaccine regimens[J]. *Sci Rep*, 2020, 10(1): 2093. DOI: 10.1038/s41598-020-57491-z.
- [21] Zhou P, Wang H, Fang M, et al. Broadly resistant HIV-1 against CD4-binding site neutralizing antibodies[J]. *PLoS Pathog*, 2019, 15(6): e1007819. DOI: 10.1371/journal.ppat.1007819.
- [22] 陈格格, 邓婷婷, 李少伟, 等. HIV-1 Env 免疫原设计的研究进展[J]. *中国艾滋病性病*, 2022, 28(5): 603-607. DOI: 10.13419/j.cnki.aids.2022.05.24.
- [23] Zolla-Pazner S, Alvarez R, Kong XP, et al. Vaccine-induced V1V2-specific antibodies control and or protect against infection with HIV, SIV and SHIV[J]. *Curr Opin HIV AIDS*, 2019, 14(4): 309-317. DOI: 10.1097/COH.0000000000000551.

(收稿日期:2024-02-28)