

技术与方法

DOI: 10.3969/j.issn.0253-3626.2012.01.019

登革病毒中国株 E 基因重组的生物信息学分析

涂 增¹, 刘利成², 白志军³, 方美玉³, 陈唯军²(1. 重庆医科大学基础医学院病原生物学教研室, 重庆 400016; 2. 中国科学院北京基因组研究所, 北京 101300;
3. 广州医学研究所, 广州 510507)

【摘要】目的: 了解我国登革病毒(Dengue virus, DV)包膜蛋白基因(Envelope, E)重组的发生情况。方法: 从 GenBank 收集共 70 个 DV 中国株的 E 基因序列, Clustal-W 比对后, 重组检测软件包 RDP3 检测 E 基因中可能存在的重组事件, 并经 Mega4 构建系统发生树验证所检测的重组事件的可靠性和真实性。结果: DV-1 型中 E 基因区存在 3 个重组事件, 共 4 个重组子, 分别是 1995、1997、1999 年和 2004 年分离的 DV 病毒株, 在其它 3 个血清型的 DV 病毒株中没有检测到重组, 也未检测到血清型间重组。结论: 我国 DV 1 型血清型 E 基因区域存在基因重组。

【关键词】RDP3; 登革热; E 基因; 遗传重组

【中国图书分类法分类号】R373.3

【文献标志码】A

【收稿日期】2011-11-30

Bioinformatics analysis of envelope gene recombination of Chinese dengue virus isolates

TU Zeng¹, LIU Licheng², BAI Zhijun³, FANG Meiyu³, CHEN Weijun²

(1. Department of Pathogen Biology, College of Basic Medical Sciences, Chongqing Medical University; 2. Beijing Genomics Institute, Chinese Academy of Sciences; 3. Guangzhou Medical Research Institute)

【Abstract】Objective: To explore the recombination of envelope protein (E) gene of Chinese dengue virus (DV) isolates. Methods: 70 Chinese DV E gene sequences available from GenBank were collected, and aligned with Clustal-W, then were incorporated into recombination detection program (RDP3) for potential recombination event detection. Phylogenetic trees constructed by Mega4 were further generated to confirm the recombination events. Results: Three intra-serotype recombination events, four recombinants isolated in 1995, 1997, 1999, and 2004, respectively, were confirmed in DV-1 E gene data. However, no recombination event was detected in other three serotypes in the data, so was the inter-serotype recombination. Conclusion: Recombination of E gene regions of Chinese dengue virus type 1 exists, although limited data were available. It highlights the necessity of employing more gene regions and sequences in the analysis to fully detect the recombination in China.

【Key words】RDP3; dengue fever; envelope protein gene; genetic recombination

登革病毒(Dengue virus, DV)属于黄病毒科黄病毒属成员, 为单股正链 RNA 病毒, 共划分成 4 个血清型(DV-1~DV-4), 借助蚊虫传播时可以引起登革热、登革出血热/登革热休克综合征等严重的热带、亚热带疾病。据 WHO 估计, 每年大约有 5 000 万到 1 亿人受到感染, 25 000 人死亡^[1,2]。我国自 1978 年首次登革热暴发到最近一次 2006 年大流行期间, 登革热每隔几年就要在南方暴发, 目前已经成为我国南方地区严重的公共健康威胁^[1,3-5]。然而, 由于尚未阐明 DV 的复制、

感染和致病机理, 迄今为止还没有临床可用的疫苗。

点突变和基因重组推动了 RNA 病毒的进化^[6]。不同血清型或基因型病毒株感染同一宿主细胞, 在其复制过程中可能发生基因片段的交换, 引起基因重组。DV 具有广泛的基因型多样性与分布的全球性^[7], 目前在不同血清型或基因型的 DV 已发现有基因重组现象^[8-11]。DV 的 E 基因全长 1 485 bp, 编码一个 54.5 kD 的糖蛋白。该蛋白包含 3 个结构域(Domain), Domain I 为 E 蛋白中心结构区, Domain II 为融合结构域, 催化囊膜与细胞膜的融合, 而 Domain III 为 IgG 样结构域, 与受体识别有关, 其在病毒免疫原性及种属特异性方面有重要作用^[1]。因此, E 基因区域的重组对病毒的适应性及进化具有潜在意义。基于此, 我们收集 GenBank 中已登陆的 DV 中国株的 E 蛋白基因, 探讨该 E 基因的重组现象。

作者介绍: 涂 增(1980-), 男, 博士,

研究方向: 病原微生物的诊断及基因组进化。

通信作者: 陈唯军, 男, 研究员, Email: chenwj@genomic.org.cn。

基金项目: 重庆医科大学青年教师招聘基金资助项目(编号: 0800280002)。

1 材料与方法

1.1 数据收集与整理

收集 GenBank 已登录的 DV 中国株 E 蛋白基因序列。Clustal-W 比对后,用重组检测软件包 RDP3 检测重组^[12],并通过 Mega4 软件以 Kimura 双参数模型构建邻接系统发生树以展示收集到的中国株间进化关系^[13]。

1.2 重组分析与确认

重组检测软件 RDP3(可从网站 <http://darwin.uvigo.es/rdp/rdp.html> 免费获取)筛选重组信号。使用 RDP3 中集成的 8 种重组检测方法 Bootscan、Chimaera、GENECONV、LARD、Max-Chi、RDP、3Seq 和 SiScan 来检测重组事件^[12]。为提高检测可信度,将 P 值设定为 0.01,且经至少 2 种方法检测到的重组事件才视为潜在重组事件。其它参数采用各软件自身的默认设置。为进一步确定重组事件的发生及可能存在的重组断

点,利用 Mega4 软件分别构建基于非重组区与重组区的邻接系统发生树^[13]。1 000 次的自举重复来验证系统发生树的可靠性。

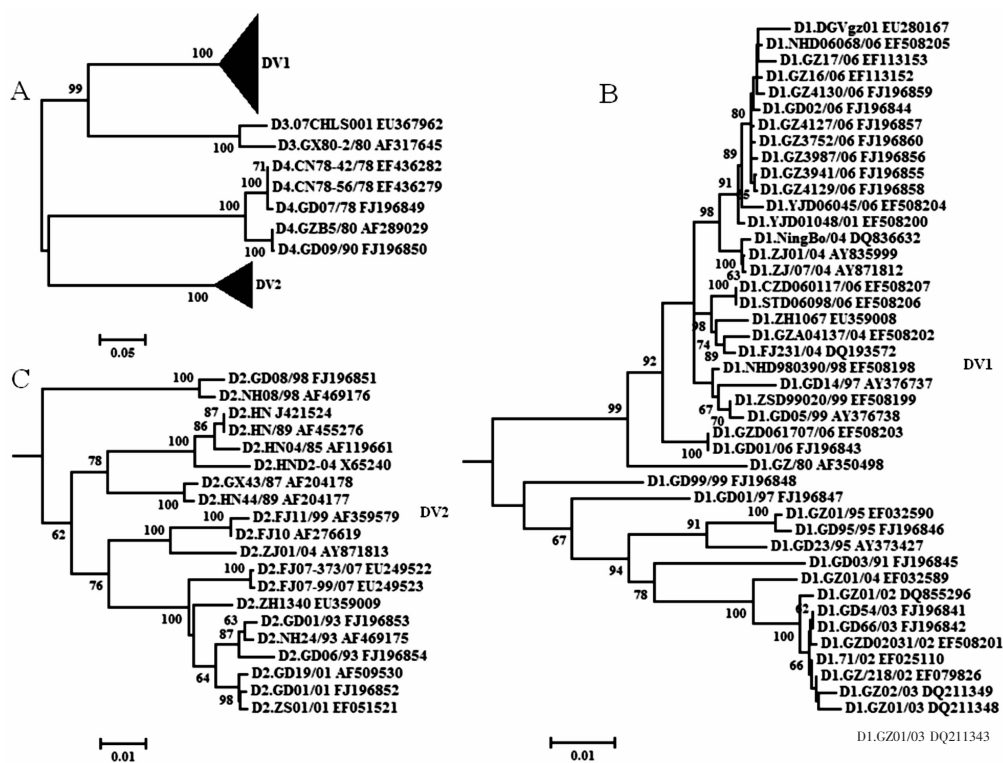
2 结果

2.1 序列收集及比对分析

总共收集到 70 个 DV 中国株的 E 基因序列(43 个 E 基因序列来自 DV-1,20 个 E 基因序列来自 DV-2,2 个 E 基因序列来自 DV-3,5 个 E 基因序列来自 DV-4)。Mega4 构建的邻接系统发生树见图 1。

2.2 重组分析

RDP3 软件包的 8 种方法共检测到 3 个重组事件,4 个重组子(表 1)。这 4 个重组子均发生于 DV-1 的 E 基因区。其它 3 种血清型登革病毒(DV-2、-3、-4)的 E 基因区域没有检测到重组,同时也没有检测到各血清型之间的重组。其中,



注:病毒株缩写:血清型/毒株编号/分离年份/GenBank 登录号;节点上的数字显示 1 000 次自举重复后节点出现的百分率,仅显示大于 60% 的节点值。分支长度与核苷酸变化成正比。A. DV 中国株 4 种血清型(DV-1, -2, -3, -4)精简的系统发生树;B、C,为 DV1 和 DV2 扩展的系统发生树

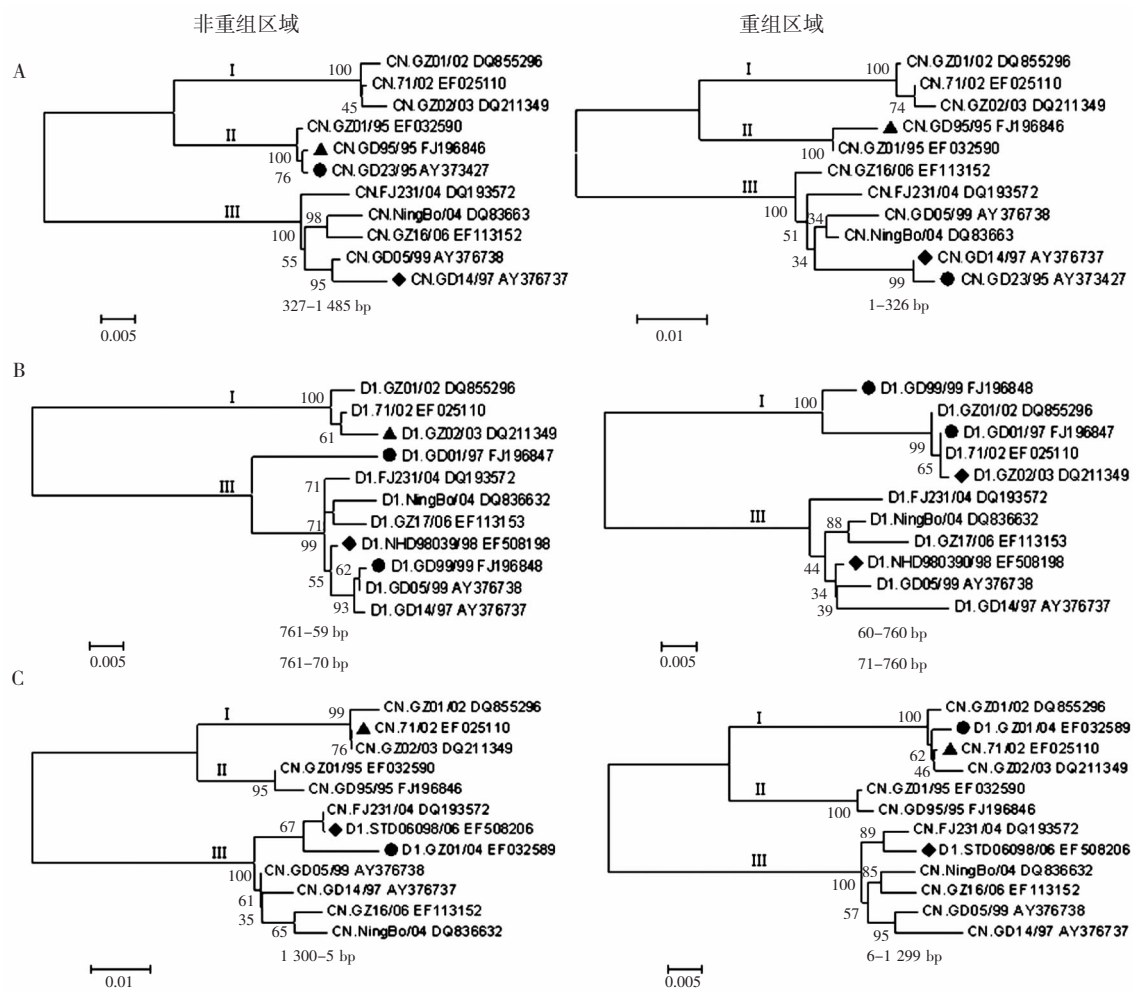
图 1 基于 70 个 DV 中国株全长 E 基因序列的邻接法系统发生树

Fig.1 Neighbor-joining (NJ) phylogenetic trees based on 70 full-length Chinese dengue E gene sequences

表 1 RDP3 检测到的 E 基因区域的重组事件

Tab.1 Recombination events of E gene detected by RDP3

重组事件	重组子	主母序列 × 次母序列	重组区域	检测方法
1	D1.GD23/95	D1.GD14/97 × D1.GD95/95	1~326	RDP, GENECONV
2	D1.GD01/97	D1.NHD98039/98 × D1.GZ02/03	60~60	BootScan, MaxChi
	D1.GD99/99		71~760	Chimaera, SiScan, LARD
3	D1.GZ01/04	D1.STD06098 × D1.71/02	6~1 299	3Seq



注: 数字表示构建系统树所用的 E 基因区域 (重组区域与表 1 相对应, 非重组区域为 E 基因重组区域之外的其它部分)。●, ◆, 和 ▲ 分别代表重组子序列、主母序列和次母序列

图 2 基于表 1 中重组事件 1 (A)、2 (B)、3 (C) 非重组区域 (左) 与重组区域 (右) Mega4 构建的邻接系统发生树

Fig. 2 NJ phylogenetic tree based on non-recombination regions (Left) and recombination regions (Right) of event 1 (A), 2 (B), and 3 (C) by Mega4

病毒株 D1.GD23/95 (AY373427) 的 E 基因的起始区 1~326 bp 为重组区域, 与 Chen 等先前报道 prM/E 基因区存在重组的情况相一致^[10], 除此之外, 其它重组事件均为首次报道。

2.3 系统发生分析

为进一步确认重组, 我们选择代表收集序列遗传多样性的病毒株序列, 分别构建 3 个重组事件基于非重组区域与重组区域 (与表 1 对应) 的系统发生树 (图 2A、B 和 C)。以重组子 D1.GD23/95 (AY373427) 为例, 在进化树中, 其非重组区域 (E 基因区 237~1 485 bp) 与 I 类次母序列 D1.GD95/95 (FJ196844) 聚类, 而其重组区域 (1~236 bp) 与 III 类主母序列 D1.GD14/97 (AY376737) 亲缘关系最近 (图 2A), 这显示 D1.GD23/95 (AY373427) 的不同 E 基因区域在进化树上位置明显不一致, 说明这 2 个区域具有不同的进化来源, 也即 D1.GD23/95 (AY373427) 为重组子, E 基因区 (1~326 bp) 为重组区域, 验证了重组的真实性。其它 2 个重组事件基于非重组区域与重组区域 (图 2B、C) 进化树同样产生类似的进化拓

扑结构变动。这些进化树的拓扑差异进一步证实了检测到的重组序列的真实性。

3 讨论

本文收集 GenBank 中收录的我国 4 个血清型的 DV 共 70 个 E 蛋白基因序列, 经重组分析和系统发生验证发现在我国的登革病毒存在重组现象, 3 个重组事件均存在于 DV-1 的 E 基因区域, 其它 3 种血清型 DV 中未检测到重组, 也未检测到不同血清型 DV 间的重组。

基因重组是病毒进化的驱动力, 目前发现在 HIV、SARS-CoV、HBV 和 HCV 等重要病毒中都存在重组现象。登革病毒重组研究较晚。1999 年 Worobey 等首先证明在登革病毒的 4 种血清型的 prM、E 和 NS1 基因中存在重组^[11,14]。而后, 随着研究人员开始关注, 并通过重组检测或系统发育分析找到更多的重组^[8-10,15,16]。Chen 等通过全基因组研究发现我国 DV 株

D1.GD23/95 (AY373427) 中存在 3 处重组, 分别位于 prM/E、NS1、NS3 区域中, 其中存在于 NS3 基因区内的重组为国内外首次报道^[10]。本文在 D1.GD23/95 (AY373427) 的 E 基因内检测到一个重组事件, 与 Chen 等的结果一致^[10]。此外, 本研究在 E 基因区域检测 2 个新的重组事件, 3 个重组子。研究证明所有 4 种血清型 DV 中均存在重组现象^[11], 但本研究没有在其他血清型 DV 中国株的 E 基因区域检测到重组, 可能原因之一是 DV 中国株序列资料本身相对较少。这些结果提示序列数据的多寡对重组检测结果产生较大影响, 因此, 全基因组分析和采用尽可能多的病毒株样本序列有助于全面检测 DV 基因重组的发生。另外, 值得注意的是, 实验室内不同病毒样品交叉污染, 在随后 PCR 扩增中也可能引起基因片段的交换, 本文获得的重组子序列与母序列分别来源于不同的实验室, 样品交叉污染可能性较低。DV 的 4 种血清型毒株常在共存于同一地理区域内, 导致人群感染一种至多种血清型病毒株; 单一血清型中不同基因型和同一基因型不同遗传背景的病毒株也可能同时流行; 媒介蚊虫常多次吸血, 常携带不同血清型或基因型病毒。上述这些情况均会增加不同血清型间或同一血清型内的 DV 毒株基因发生重组的机会。E 基因编码 DV 的包膜蛋白, 决定该病毒的免疫性原性和细胞嗜性。目前已报道的重组断点多位于编码 E 蛋白结构域 II 融合多肽的核苷酸序列区域内^[10,11,14-16], 此区域可能是 DV 基因重组的热点, 但这一区域内基因重组的分子基础及生物学意义亟待进一步研究。

基因重组影响病毒分子进化、分子诊断及疫苗开发等研究。E 基因的重组可能在对 DV 分型时引起误判。目前常用的 DV 分子分型方法是利用全长或部分 E 基因或 E/NS1 连接区域构建进化树而将 DV 分型, 而潜在的 E 基因重组常被忽略, 从而可能造成分型误差。同样, 对分子诊断而言, 诊断靶标的选择不仅要考虑区域的保守或特异性, 同时也要考虑到无重组发生的区域, 这样诊断才可能更准确。另外, 由于不同 DV 血清型间存在抗体依赖性增强现象, 故而 4 价 DV 疫苗才能满足免疫需要, 但这些弱化的疫苗株会不会因基因重组而导致免疫丢失或毒力增强等生物学性质的改变有待研究。所以, 更广泛、更深入地探讨 DV 遗传重组、其分子机制及对病毒生物学影响, 有助于更准确地理解病毒的分子进化并有助于病毒疾病的预防与治疗。

参 考 文 献

- [1] 秦鄂德, 秦成峰, 姜涛. 登革病毒和登革病毒病[M]. 北京: 科学出版社, 2008.212.
- Qin E D, Qin C F, Jiang T. Dengue virus and dengue viral disease[M].

Beijing: Science Press, 2008.212.

- [2] Guzman M G, Halstead S B, Artsob H, et al. Dengue: a continuing global threat[J]. Nat Rev Microbiol, 2010, 8(12 s): 7-16.
- [3] Chen S P. The origin of dengue viruses caused the DF outbreak in Guangdong province, China, in 2006[J]. Infect Genet Evol, 2011, 11(5): 1183-1187.
- [4] Wu J Y, Lun Z R, James A A, et al. Dengue fever in mainland China[J]. Am J Trop Med Hyg, 2010, 83(3): 664-671.
- [5] 张复春. 登革热: 一个日益严重的全球性公共卫生问题[J]. 实用医学杂志, 2011, 27(19): 3459-3461.
- Zhang F C. Dengue: an increasing public health problem[J]. J Prac Med, 2011, 27(19): 3459-3461.
- [6] Barr J N, Fearn R. How RNA viruses maintain their genome integrity[J]. J General Virol, 2010, 91(6): 1373-1378.
- [7] Holmes E C, Twiddy S S. The origin, emergence and evolutionary genetics of dengue virus[J]. Infect Genet Evol, 2003, 3(1): 19-28.
- [8] Aaskov J, Buzacott K, Field E, et al. Multiple recombinant dengue type 1 viruses in an isolate from a dengue patient[J]. J Gen Virol, 2007, 88(12): 3334-3340.
- [9] Perez-Ramirez G, Diaz-Badillo A, Camacho-Nuez M. Multiple recombinants in two dengue virus, serotype-2 isolates from patients from Oaxaca, Mexico[J]. BMC Microbiol, 2009, 9: 260.
- [10] Chen S P, Yu M, Jiang T, et al. Identification of a recombinant dengue virus type 1 with 3 recombination regions in natural populations in Guangdong province, China[J]. Arch Virol, 2008, 153(6): 1175-1179.
- [11] Worobey M, Rambaut A, Holmes E C. Widespread intra-serotype recombination in natural populations of dengue virus[J]. Proc Natl Acad Sci USA, 1999, 96(13): 7352-7357.
- [12] Martin D P, Lemey P, Lott M, et al. RDP3: a flexible and fast computer program for analyzing recombination[J]. Bioinformatics, 2010, 26(19): 2462-2463.
- [13] Tamura K, Dudley J, Nei M, et al. Mega4: Molecular evolutionary genetics analysis (MEGA) software version 4.0[J]. Mol Biol Evol, 2007, 24(8): 1596-1599.
- [14] Holmes E C, Worobey M, Rambaut A. Phylogenetic evidence for recombination in dengue virus[J]. Mol Biol Evol, 1999, 16(3): 405-409.
- [15] Tolou H J G, Couissinier-Paris P, Durand J P, et al. Evidence for recombination in natural populations of dengue virus type 1 based on the analysis of complete genome sequences[J]. J Gen Virol, 2001, 82(6): 1283-1290.
- [16] Uzcategui N Y, Camacho D, Comach G, et al. Molecular epidemiology of dengue type 2 virus in Venezuela: Evidence for in situ virus evolution and recombination[J]. J Gen Virol, 2001, 82(12): 2945-2953.

(责任编辑: 张学颖)