Infection, Genetics and Evolution 15 (2013) 77–86

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Complex dynamic of dengue virus serotypes 2 and 3 in Cambodia following series of climate disasters

Veasna Duong [a], Matthew R. Henn [b], Cameron Simmons [c], Chantha Ngan [d], Bunthin Y [a], Laurent Gavotte [e], Alain Viari [f], Sivuth Ong [a], Rekol Huy [d], Niall J. Lennon [b], Sowath Ly [a], Sirenda Vong [a], Bruce W. Birren [f], Jeremy J. Farrar [c], Vincent Deubel [a], Roger Frutos [g], Philippe Buchy [a,*]

[a] Institut Pasteur in Cambodia, Réseau International des Instituts Pasteur, 5 Monivong Boulevard, P.O. Box 983, Phnom Penh, Cambodia
[b] Broad Institute, Cambridge, MA, USA
[c] Hospital for Tropical Diseases, Oxford University Clinical Research Unit, Wellcome Trust Major Overseas Programme, Ho Chi Minh City, Viet Nam
[d] National Dengue Control Program, National Center of Parasitology, Entomology and Malaria Control, Ministry of Health, Cambodia
[e] UM2, ISEM, UMR 5554, CNRS-UM2-IRD, Université Montpellier 2 CC065, Place E. Bataillon, 34095 Montpellier Cedex 5, France
[f] INRIA Grenoble Rhône-Alpes, Projet Bamboo, 655 Av. de l'Europe, 38330 Montbonnot-Saint Martin, France
[g] CIRAD, UMR17, CIRAD-IRD, TA-A17/G, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France

ARTICLE INFO

Article history:
Available online 5 June 2012

Keywords:
Dengue virus
DENV-2
DENV-3
Diversity
Lineage dynamics
Cambodia

ABSTRACT

The Dengue National Control Program was established in Cambodia in 2000 and has reported between 10,000 and 40,000 dengue cases per year with a case fatality rate ranging from 0.7 to 1.7. In this study 39 DENV-2 and 57 DENV-3 viruses isolated from patients between 2000 and 2008 were fully sequenced. Five DENV2 and four DENV3 distinct lineages with different dynamics were identified. Each lineage was characterized by the presence of specific mutations with no evidence of recombination. In both DENV-2 and DENV-3 the lineages present prior to 2003 were replaced after that date by unrelated lineages. After 2003, DENV-2 lineages D2–3 and D2–4 cocirculated until 2007 when they were almost completely replaced by a lineage D2–5 which emerged from D2–3 Conversely, all DENV-3 lineages remained, diversified and cocirculated with novel lineages emerging. Years 2006 and 2007 were marked by a high prevalence of DENV-3 and 2007 with a large dengue outbreak and a high proportion of patients with severe disease. Selective sweeps in DENV-1 and DENV-2 were linked to immunological escape to a predominately DENV-3-driven immunological response. The complex dynamic of dengue in Cambodia in the last ten years has been associated with a combination of stochastic climatic events, cocirculation, coevolution, adaptation to different vector populations, and with the human population immunological landscape.

1. Introduction

Dengue is the most rapidly spreading mosquito-borne viral disease with a 30-fold rise in the number of human cases reported in the last 50 years associated with a geographic expansion of the disease to new countries (WHO/TDR, 2009). Globally, about two-thirds of the world's population lives in area at risk for infection (more than 75% in the WHO South-East Asia and Western Pacific regions) and an estimated 50–100 million cases of dengue infection occur every year (WHO/TDR, 2009). Dengue has a wide spectrum of clinical presentations, often with unpredictable clinical evolution and outcome. While most patients recover following a self-limiting non-severe clinical course, a small proportion progress to severe disease, mostly characterized by plasma leakage with or without hemorrhage (Duong et al., 2009; WHO/TDR, 2009).

Dengue virus (DENV) which comprises four genetically and antigenically distinct serotypes (DENV-1, -2, -3, and -4) belongs to the family *Flaviviridae*, genus *Flavivirus* (Calisher et al., 1989). DENV genome is a single stranded positive-sense RNA virus and has about 11 kb in length encoding a single open reading frame (Lindenbach and Rice, 2003). The translated protein is cleaved by host- and virus-derived proteases to produce structural proteins (capsid, premembrane/membrane, envelope; C, prM/M, E) and non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) (Lindenbach and Rice, 2003; Weaver and Vasilakis, 2009). The single open reading frame (ORF) is flanked by 5′ untranslated region (UTR) capped with type I 7-methyl guanosine structure and by 3′ UTR lacked of poly(A) region (Lindenbach and Rice, 2003). Nucleic acid sequencing has allowed the classification of each of the DENV serotype into genotypes (Rico-Hesse, 1990). Rico-Hesse defined these genotypes as clusters of DENV viruses

* Corresponding author. Address: Virology Unit, Institut Pasteur in Cambodia, 5 Monivong Boulevard, P.O. Box 983, Phnom Penh, Cambodia. Tel.: +855 12802982.
E-mail address: pbuchy@pasteur-kh.org (P. Buchy).

1567-1348/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2012.05.012

having nucleotide sequence divergence not greater than 6% within a given genome region (in this case the E/NS1 junction). Various phylogenetic analyses based on partial E/NS1 or complete E nucleotide sequences indicated that (1) DENV-1 are organized into five genotypes (I–V), (2) DENV-2 comprises six genotypes: South-East Asian/American, Asian I, Asian II, Cosmopolitan, American and sylvatic, (3) DENV-3 comprises of 4 genotypes (I–IV) and (4) DENV-4 are classed into 4 genotypes (I, II, III and sylvatic) (Holmes and Twiddy, 2003; Rico-Hesse, 2003; Vasilakis and Weaver, 2008; Weaver and Vasilakis, 2009).

Previous studies have shown that dengue genotypes are not fixed entities (Holmes et al., 1999; Tolou et al., 2001; Worobey et al., 1999) and evidence of selection pressure showing lineage turnover (Sittisombut et al., 1997; Vu et al., 2010; Wittke et al., 2002; Zhang et al., 2005). Although, the mechanism of lineage replacement's occurrence is unclear, two main hypothesis have been proposed (Zhang et al., 2005): (1) a ladder-like phylogenetic trees showing a strong temporal topology. This event may be introduced by the elimination of deleterious mutation strains by purifying selection (Holmes, 2003) and/or a regular random population bottleneck perhaps due to decline in mosquito population and density during the annual dry season (Scott et al., 2000); (2) or a more dramatic change by entire clade replacement by a new clade of viruses. The virus evolutionary process remains unclear but the patterns observed may reflect the action of either dramatic population bottlenecks or natural selection such as clades with an improved fitness which out-compete previously circulating clades or lineages of viruses with mutations that allow them to evade cross-protective herd immunity (Sittisombut et al., 1997; Vu et al., 2010; Wittke et al., 2002). In Cambodia, all the four DENV serotypes co-circulate each year although the predominant serotype has alternated mainly between DENV-2 and DENV-3 during the last decade. Therefore, separate in depth analyses of the lineage structure and dynamics of each serotype were undertaken over the period 2000–2008. We report in this study the first of this series of analyses: the characterization of the genetic diversity and lineage dynamic of the complete genome of DENV-2 and DENV-3 viruses isolated in Cambodia between 2000 and 2008.

2. Material and methods

2.1. Virus strains

Dengue virus strains were isolated by the national dengue laboratory at Institut Pasteur in Cambodia established to support the five sentinel sites from the National Dengue Control Program (NDCP), Ministry of Health Cambodia, and by a dengue cohort study in Kampong Cham province (Vong et al., 2010). This study received approval from the National Ethics Committee for Health Research in Cambodia. Written informed consent was given by all patients (or their parents/guardians) included in the cohort study. Samples collection methods and diagnostic testing including serological, molecular and virological tests were previously described (Buchy et al., 2005; Huy et al., 2010). All strains were isolated in C6/36 mosquito cells (*Aedes albopictus* clone) and cultures were not passed more than two times. After anonymization, the strains included in the study were selected to be representative of geographical origin, year of sampling and severity of the disease according to former WHO criteria (WHO, 1997).

2.2. DENV genome sequencing

Viral genomes were sequenced using the Broad Institute's capillary sequencing (Applied Biosystems) directed amplification viral sequencing pipeline (http://www.broadinstitute.org/annotation/viral/Dengue; (Vu et al., 2010). Briefly, viral RNA was extracted from low passage cell culture supernatants using the QIAmp viral RNA mini kit (Qiagen). The genome was reverse-transcribed to cDNA with SuperScript III reverse transcriptase (Invitrogen, California, USA), random hexamers (Roche Diagnostics GmbH, Mannheim, Germany) and a specific oligonucleotide targeting the 3′ end of the target genome sequences (5′-AGA ACC TGT TGA TTC AAC AGC AC-3′; nt 10700–10722). The cDNA obtained was then amplified using a high fidelity DNA polymerase (Roche Diagnostics GmbH, Mannheim, Germany) and a pool of specific primers to produce 14 overlapping amplicons of 1.5–2 kb in size for a physical coverage of 2-fold. Amplicons were then sequenced in the forward and reverse direction using primer panels consisting of 96 specific primer pairs, tailed with M13 forward and reverse primer sequence, that produce 500–700 bp amplicons from the target viral genome. Total sequence coverage delivered post amplification and sequencing was about 8-fold. Resulting sequence reads were assembled *de novo* and annotated using the Broad Institute's AV454 algorithm (Henn et al., in press) and an in-house annotation algorithm.

2.3. Sequence alignment and phylogenetic analysis

Multiple sequence alignment was conducted with Muscle (Edgar, 2004) available in Seaview version 4.2.5 package (Galtier et al., 1996; Gouy et al., 2010). Phylogenetic analyses were performed using maximum likelihood (ML) method for the complete coding region and for each gene separately for both the 39 DENV-2 and 57 DENV-3 isolates. Jmodeltest (Posada and Crandall, 1998) was used to select the optimal evolution model by evaluating the selected parameters using the Akaike Information Criterion (AIC). A corrected version of the AIC (AICc) was used because the sample size (n) was small compared with the number of parameters ($n/K < 40$). This approach suggested the following models: GTR + G4 for complete coding region, HKY + I + G for C; GTR + G for E, NS1 and NS2A; GTR + I + G for NS2B, NS3, NS4A, NS4B, NS5 and concatenated data set; and K80 + I for PrM. Under the selected models, the parameters were optimized and ML analyses were performed with PhyML (version 2.4.4) (Guindon and Gascuel, 2003). The robustness of nodes was assessed with 1000 booststrap replicates for complete coding region, 100 bootstrap replicates for each gene and 500 bootstrap replicates for concatenated data set. Finally, trees were edited using FigTree v1.3.1 (BEAST software).

2.4. Genetic similarity and DNA polymorphism analysis

Genetic similarity and difference matrices were constructed from ClustalX2 alignments (Thompson et al., 1997) using BioEdit 7.0.9.0 (Hall, 1999). Phylogenetic analyses do not provide a comprehensive view of the relative dynamics and evolutionary patterns of the sequences analyzed. To assess the presence of specific selective pressure and evolutionary patterns a DNA polymorphism analysis was conducted on each separate gene for the various lineages identified. DNA sequence polymorphism and all subsequent tests were investigated using several functions from the DnaSP5.00.02 package (Librado and Rozas, 2009). Haplotypes (alleles) were calculated according to Nei (Nei, 1987). Polymorphism was assessed using a the following parameters: nucleotide diversity, Pi (π), the average number of nucleotide differences per site between two sequences was calculated according to Nei (Nei, 1987), using the Jukes and Cantor (Jukes and Cantor, 1969) correction. Eta (η) is the total number of mutations, and S is the number of segregating (polymorphic) sites. Si is the number of singletons (one unique mutation at a given site) and Pa is the number parsimony informative sites (several mutations at a given site). Theta (θ) (Watterson's mutation parameter) was calculated for the whole sequence from η (Watterson, 1975). *Na* is the number