CHRISTIAN JULIÁN VILLABONA ARENAS

**PHYLOGEOGRAPHY OF THE 2013 URBAN OUTBREAK OF DENGUE VIRUS IN GUARUJÁ, SÃO
PAULO**

Tese apresentada ao Departamento de Microbiologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção do Título de Doutor em Ciências.

São Paulo
2014

CHRISTIAN JULIÁN VILLABONA ARENAS

**PHYLOGEOGRAPHY OF THE 2013 URBAN OUTBREAK OF DENGUE VIRUS IN GUARUJÁ, SÃO
PAULO**

Tese apresentada ao Departamento de Microbiologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção do Título de Doutor em Ciências.

Área de concentração: Microbiologia

Orientador: Prof. Dr. Paolo Marinho de Andrade Zanotto

Versão original

São Paulo
2014

DADOS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Serviço de Biblioteca e Informação Biomédica do
Instituto de Ciências Biomédicas da Universidade de São Paulo

reprodução não autorizada pelo autor

Villabona Arenas, Christian Julián.

Filogeografia do surto urbano de 2013 da dengue em Guarujá, São Paulo / Christian Julián Villabona Arenas. -- São Paulo, 2014.

Orientador: Prof. Dr. Paolo Marinho Andrade Zanotto.

Tese (Doutorado) – Universidade de São Paulo. Instituto de Ciências Biomédicas. Departamento de Microbiologia. Área de concentração: Microbiologia. Linha de pesquisa: Dengue.

Versão do título para o inglês: Phylogeography of the 2013 urban outbreak of dengue virus in Guarujá, São Paulo.

1. Dengue 2. Evolução 3. Filogenia 4. Epidemiologia 5. Vírus do dengue 6. Surtos de doenças I. Zanotto, Prof. Dr. Paolo Marinho Andrade II. Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação em Microbiologia III. Título.

ICB/SBIB0175/2014

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Candidato(a): Christian Julián Villabona Arenas.

Título da Tese: Filogeografia do surto urbano de 2013 da dengue em Guarujá, São Paulo.

Orientador(a): Prof. Dr. Paolo Marinho Andrade zanotto.

A Comissão Julgadora dos trabalhos de Defesa da Tese de Doutorado, em sessão pública realizada a/...../....., considerou

() Aprovado(a)

() Reprovado(a)

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Presidente: Assinatura:
Nome:
Instituição:

Research ethics disclaimer

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS**

Cidade Universitária "Armando de Salles Oliveira"
Av. Prof. Lineu Prestes, 2415 - cep. 05508-000 São Paulo, SP - Brasil
Telefone : (55) (11) 3091.7733 telefax : (55) (11) 3091-8405
e-mail: cep@icb.usp.br

São Paulo, 26 de março de 2010.

PARECER 933 /CEP

A Comissão de Ética em Pesquisas com Seres Humanos do ICB, na sessão de 25.03.2010 **APROVOU** o projeto intitulado: "*Filo-dinâmica da dengue em focos urbanos no Município de Jundiaí, São Paulo, Brasil*" sob responsabilidade de execução dos autores Prof. Dr. **PAOLO MARINHO ANDRADE ZANOTTO** e o aluno **CHRISTIAN JULIÁN VILLABONA-ARENAS**.

Cabe aos pesquisadores executantes elaborar e apresentar a este Comitê, relatórios anuais (parciais ou final), de acordo com a resolução 196/06 do Conselho Nacional da Saúde, item IX.2 letra c.

O primeiro relatório deverá ser encaminhado à Secretaria deste CEP em **25.03.2011**.

Atenciosamente,

Prof. Dr. PAOLO MARINHO ANDRADE ZANOTTO
Vice-Coordenador da Comissão de Ética em
Pesquisas com Seres Humanos - ICB/USP

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Cidade Universitária "Armando de Salles Oliveira"
Av. Prof. Lineu Prestes, 2415 - cep. 05508-000 São Paulo, SP - Brasil
Telefone: (55) (011) 3091.7733 - e-mail: cep@icb.usp.br

Decl. CEP SH. 016 /13.

DECLARAÇÃO

Em adendo ao PARECER 933/CEP, de 26.03.2010, informo que está Comissão em reunião de 27.02.13 aprovou a alteração do título do Projeto para "*Filogeografia dos vírus da dengue nos municípios de Jundiaí e Guarujá no estado de São Paulo*" sob responsabilidade dos autores Prof. Dr. **PAOLO MARINHO ANDRADE ZANOTTO** e o aluno **CHRISTIAN JULIÁN VILLABONA-ARENAS**, não apresentando restrições quanto às modificações do projeto, que as mesmas não afetam os aspectos éticos da pesquisa.

São Paulo, 01 de março de 2013.

Prof. Dr. Paolo M.A. Zanotto
Coordenador da Comissão de Ética em
Pesquisas com Seres Humanos - ICB /USP

For the women in my life: my mother and sisters

Elizabeth, Nathalia and Stephany.

For my Happy New Year, Gui.

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisor, Dr. Paolo Marinho de Andrade Zanotto, whose expertise, added considerably to my graduate experience.

I would like to thank the other members (and ex-members) of my lab, Camila Romano, Carla Braconi, Danielle Ferreira, Juliana Bloch, Juliana Velasco, Anderson Brito, Atila Iamarino, Caio Freire, Fernando Melo, Nicholas Di Paola, Susana Ortiz, Victor Pimentel for the help they provided at all levels of my instruction. Special thanks go for Jessica Oliveira whose skills and assistance were fundamental to complete the research project.

I would also like to thank my family for the support they provided me through my entire life. I would also like to thank my friends Adriana Torres, Diana Martinez, Kattia Palacio, Natalia Sierra, Oriana Nader, Renata Gomes, Viviane Casaroli, Vanesa Castro, Alexandre Suenaga, Bernardo Machado, Braulio Carli, Chalo Gonzalez, Felix Martinez, Francisco Gomez, Gabriel Vasques, Renato Braghiere, Victor Bastos and Vagner Andrade for our philosophical debates, exchanges of knowledge, skills, and venting of frustration during my graduate program, which helped enrich the experience. I must acknowledge Guillermo, whose encouragement, love and patience kept me going.

I would like to thank Carla Cappa from the Guarujá Municipal Health Department for taking time out from his busy schedule to help me with the sampling. It was through her persistence and consideration that the research lifted off.

Appreciation also goes out to the professors for all of assistance throughout my graduate program, and to the Department staff for all the instances in which their assistance helped me along the way. An especial thanks goes to Gisele da Graça Santana and to those who provided me with advice at different times: Teresa Furlan and Saulos Passos.

Finally, I recognize that this research would not have been possible without the financial assistance of CAPES, FAPESP (Grants 2010/19059-7 and 2011/17071-2), the Guarujá Municipal Health Department, and Itapema Medical Laboratory and express my gratitude to those institutions.

“You can tell the story of the Broad Street outbreak on the scale of a few hundred human lives, people drinking water from a pump, getting sick and dying over a few weeks, but in telling the story that way, you limit its perspective, limits its ability to convey a fair account of what really happened, and more importantly -*why* it happened. Once you get to why, the story has to widen and tighten at the same time: to the long *durée* of, urban development or the microscopic tight focus of bacterial life cycles. These are causes, too”.

Steven Johnson

ABSTRACT

VILLABONA-ARENAS, C. J. **Phylogeography of the 2013 urban outbreak of dengue virus in Guarujá, São Paulo**. 2014. 77 p. Ph. D. thesis (Microbiology) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2014.

Dengue virus type 1 (DENV-1) was introduced in Brazil in 1986 and caused several epidemics. The first autochthonous cases of DENV-2 and DENV-3 were detected respectively in 1990 and 2000. Since then, the viruses have spread throughout Brazil and became endemic in most areas infested with *Aedes aegypti*. DENV-4 was isolated for the first time in 1982 in a focal epidemic in the northwestern region of the Brazilian Amazon. Later, in 2008, this serotype emerged as an important pathogen during outbreaks. The study of the historical processes that may be responsible for the contemporary geographic distributions of viruses is critical to understand viral epidemiology. However, those processes in urban scales are not well understood. 2013 was one of the worst years for dengue in the Brazil's history, with 1.4 million cases, including 6,969 severe cases and 545 deaths. This project aimed to understand the dynamics of evolutionary change, origins and distributions of different viral strains in an urban setting during 2013. We expect this study to provide new perspectives for viral control.

Keywords: Dengue Virus. Epidemiology. Genetic Diversity. Molecular Evolution. Phylogeny. Phylogeography.

RESUMO

VILLABONA-ARENAS, C. J. **Filogeografia do surto urbano de 2013 da Dengue em Guarujá, São Paulo**. 2014. 77 f. Tese (Doutorado em Microbiologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2014.

O vírus da dengue tipo 1 (DENV-1) foi introduzido no Brasil em 1986 e foi responsável por numerosas epidemias. Os primeiros casos autóctones do DENV-2 e DENV-3 foram detectados respectivamente em 1990 e 2000. Desde então, o vírus tem se espalhado por todo o Brasil e tornou-se endêmico na maioria das áreas infestadas com *Aedes aegypti*. DENV-4 foi isolado pela primeira vez em 1982, em uma epidemia focal na região noroeste da Amazônia brasileira. Porém, este sorotipo somente emergiu como um importante patógeno durante os surtos de 2008. O estudo dos processos históricos que podem ser responsáveis para as distribuições geográficas contemporâneas do vírus é fundamental para compreender a epidemiologia viral. No entanto, esses processos em escalas urbanas não são bem compreendidos. 2013 foi um dos piores anos para a dengue na história do Brasil, com 1,4 milhões de casos, incluindo 6.969 casos graves e 545 mortes. Este projeto teve como objetivo compreender a dinâmica de mudança evolutiva, origens e distribuições de diferentes cepas virais em um cenário urbano em 2013. Esperamos que este estudo contribua com novas perspectivas para o controle viral.

Palavras-chave: Diversidade genética. Epidemiologia. Evolução molecular. Filogenia. Filogeografia. Vírus da Dengue.

TABLE OF CONTENTS

1 INTRODUCTION.....	13
1.1 Flavivirus in brief.....	13
1.2 Dengue viruses.....	15
1.3 Dengue in numbers.....	18
1.4 Dengue in Brazil.....	20
1.4.1 History.....	20
1.4.2 Dengue in the State of São Paulo.....	21
1.4.3 Reinfestation in Brazil and in the State of São Paulo.....	22
1.5 Urban Dengue Outbreaks.....	23
2 OBJECTIVES.....	25
3 METHODS.....	27
3.1 Sample collection.....	27
3.2 Molecular testing.....	29
3.3 Genetic analysis.....	31
3.4 Phylogenetic reconstruction.....	31
3.5 Phylodynamics.....	33
3.6 Spatio-temporal dispersion pattern.....	35
3.7 Transmission tree.....	35
3.8 Extrinsic traits.....	36
4 RESULTS AND DISCUSSION.....	37
4.1 Sampling.....	37
4.2 Hyperendemicity of dengue.....	39
4.3 DENV-4 sequence analysis.....	42
4.4 Evolutionary relationships of DENV-4.....	43
4.5 Epidemiological findings.....	46
4.6 Spread dynamics.....	50
4.7 Transmission dynamics.....	53
4.7.1 Patient attributes and transmission.....	58
4.8 Filling the gap.....	60
5 CONCLUSIONS.....	65

REFERENCES*67

APPENDIX A – Manuscripts.....77

1 INTRODUCTION

1.1 Flavivirus in brief

The dengue viruses (DENV) belong to the genus *Flavivirus*, family *Flaviviridae*. This genus comprises more than 70 different viruses. Many of these viruses are transmitted by either mosquitoes or ticks and for that reason they are also classified as arboviruses (arthropod-borne-viruses) (KARABATSOS, 1985). Kuno et al. (1998) showed that phylogenetically the members of the genus fall into two major branches: non-vector and vector-borne clades; the latter diverged into the tick-borne and mosquito-borne groups (See Figure 1). Different clades correlate with previous classifications based on antigenic complexes.

All flaviviruses are enveloped viruses with a single stranded positive-sense RNA genome of approximately 11 kb long. The genome encodes a single open reading frame (ORF), flanked by highly structured 5' and 3' untranslated regions (UTRs).

Most human infections with flaviviruses are typically incidental – man is a dead end host. Nonetheless, some viruses have established primary life cycle exclusively involving transmission between humans and vectors.

Figure 1 - Phylogenetic relationships among Flaviviridae viruses.

Flavivirus genus is highlighted in blue, Pestivirus is in green and Hepacivirus is in salmon. Flavivirus genus was further divided to reflect distinct viral host range, with insect-only viruses in orange, not known vector in purple, mosquito-borne in yellow and tick-borne in red. Source: Lobo et al., 2009

1.2 Dengue viruses

The dengue viruses exist as four antigenically distinct serotypes named DENV-1, DENV-2, DENV-3 and DENV-4. Recent work documents a novel serotype, formerly DENV-5 (NORMILE, 2013). The term serotype is used to describe viruses that induce an overlapping immune response to each other. Infection with any DENV provides long-term protection against infection with another virus of the same serotype, but only short-lasting immunity to the other serotypes (BURKE et al., 1988; SABIN, 1952). As a consequence, secondary infections are frequent after cross-immunity has waned; tertiary and quaternary infections, although possible, are very rare (GIBBONS et al., 2007). In addition, there is evidence that secondary infections are enhanced via immune interactions (GUZMAN; ALVAREZ; HALSTEAD, 2013; HALSTEAD, 1988;).

Within each of the four serotypes there is considerable diversity, reflected in the presence of clusters of lineages, designated genotypes (HOLMES, 2004; RICO-HESSE, 1990). These genotypes often relates to the region where particular strains are commonly found or were first isolated (e.g. DENV-2 American/Asian genotype).

The virus enters the host cell by receptor-mediated endocytosis and upon internalization and acidification of the endosome the viral and vesicular membranes fused and allows release of the genomic RNA into the cytoplasm. The viral RNA (See Figure 2) serves as mRNA for translation and subsequently, as template for RNA synthesis. Replication of the viral RNA occurs in the context of complex three-dimensional networks of membranes induced by the viral non-structural proteins. Virus assembly occurs on membranes derived from the endoplasmic reticulum (ER). Virions bud into the ER as immature virus particles that incorporate the pre-membrane (prM) and envelope proteins. During egress, prM is cleaved by the cellular serine protease furin. A relatively smooth, infectious, mature virus particle is released into the extracellular space (PIERSON, 2012).

Dengue fever (DF) is a disease caused by any of the DENV. The frequent label of *breakbone fever* comes from a popular name used to describe an illness that is clinically compatible with DF and occurred in the city of San Juan (Puerto Rico) in 1771 and in the city of Philadelphia (Pennsylvania) in 1780 (RIGAU-PEREZ, 1998; RUSH, 1809).

Figure 2 - Schematic representation of the Dengue virus (DENV) genome.

This RNA encodes a single ORF, which is flanked by untranslated regions (UTR) of 96 and ~450 nucleotides, respectively. The 5' UTR has a type 1 cap (m7GpppAmp) structure and the 3' UTR lacks a poly (A) tail. Conserved structural elements in the 3' UTR are involved in viral replication, regulation of translation, and RNA synthesis, as well as in interactions with viral and cellular proteins. The genome serves as an mRNA for the translation of the viral proteins and encodes three structural proteins — capsid (C), membrane (prM; processed to M) and envelope (E) — and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5). Translation of the ORF produces a large polyprotein that is cleaved by host signal peptidases and a viral serine protease during and after translation to yield the ten viral proteins. Some of the functions that are carried out by the viral proteins are indicated in the figure. Sources: modified from Gebhard, Filomatori and Gamarnik (2011) and Vasilakis et al. (2011).

The principal symptoms of DF are high fever, severe headache, severe pain behind the eyes, joint pain, muscle and bone pain, rash, and mild bleeding. Severe dengue is a potentially deadly complication due to plasma leaking, fluid accumulation, respiratory distress, severe bleeding, or organ impairment. Generally, younger children and those with their first dengue infection have a milder illness than older children and adults (CENTERS FOR DISEASE CONTROL AND PREVENTION, 2014).

DENV are transmitted among people by the mosquitoes *Aedes aegypti* and *Aedes albopictus* (Figure 3). Symptoms of infection usually begin 4-7 days after an infected mosquito bite (referred as intrinsic incubation period, IIP) and typically last 3-10 days. In order for transmission to occur another mosquito must feed on a person in the five days period when large amounts of virus are in the blood; this period usually begins a little before the person become symptomatic. Some people never have significant symptoms but can still infect mosquitoes. After the mosquito ingests the virus in the blood meal, it requires an additional 8-12 days incubation (referred as extrinsic incubation period, EIP) before it can

then be transmitted to another human. The mosquito remains infected for the remainder of its life, which might be days or a few weeks (CDC, 2014).

Two distinct DENV transmission cycles are recognized: an endemic/epidemic cycle and a sylvatic cycle (Figure 3). The former involves the human host and *A. aegypti* and *A. albopictus* as vectors. The sylvatic transmission cycle involves monkeys and several different *Aedes* mosquitoes from Asia and West Africa. Humans can be infected with sylvatic DENV, but there have been no sustained epidemics (CHEN; VASILAKIS, 2011).

Figure 3- The Transmission Cycles of Dengue Virus

The transmission cycles of DENV, depicting the sylvatic origins and the “zone of emergence” where sylvatic cycles contact human populations in rural areas in West Africa and Southeast Asia. Uninfected hosts/vectors are colored in blue, infected host/vectors are colored in salmon. Source: modified from Whitehead et al., 2006; Chen; Johansson, 2012; Chen; Vasilakis, 2011.

The original four serotypes originated in monkeys and independently jumped to humans in Africa or Southeast Asia over a century ago (WANG et al., 2000). Dengue remained a relatively minor, geographically restricted disease until the middle of the 20th century when the disruption of the Second World War allows the transport of *Aedes* mosquitoes around the world in cargo. Hence, the mosquito expanded its range and disseminated the viruses (CDC, 2014). This situation, coupled with increased human population growth and long-distance travel, rapid urbanization, lack of sanitation and

ineffective mosquito control have resulted in sustain disease transmission (GUBLER 1998, 2006).

1.3 Dengue in numbers

Nowadays, about 2.5 billion people, or 40% of the world's population, live in tropical and subtropical areas where there is a risk of dengue transmission. The World Health Organization (WHO) (2014) estimates that 50 to 100 million infections occur yearly, including 500,000 severe cases and 22,000 deaths, mostly among children. In contrast, Bhatt et al. (2013) using cartographic approaches estimated 390 million (95% credible interval 284–528) dengue infections per year, of which 96 million (67–136) manifest apparently (any level of clinical or subclinical severity) (Figure 4).

The number of dengue cases in the Americas increased five-fold between 2003 and 2013. Between 2009 and 2012, over 1 million cases were reported annually, on average, with more than 33,900 severe cases and 835 deaths. In 2013, 2.3 million cases were reported region-wide, including 37,705 severe cases and 1,289 deaths. By comparison, the number of cases reported in 2003 was 517,617 (PAN AMERICAN HEALTH ORGANIZATION, 2014).

Figure 4 - Global evidence consensus, risk and burden of dengue in 2010.

(a) National and subnational evidence consensus on complete absence (green) through to complete presence (red) of dengue. (b) Probability of dengue occurrence. Areas with a high probability of dengue occurrence are shown in red and areas with a low probability in green. (c) Cartogram of the annual number of infections for all ages as a proportion of national or subnational (China) geographical area. Source: Bhatt et al., 2013.

Dengue is endemic in at least 100 countries in Asia, the Pacific, the Americas, Africa, and the Caribbean. This means that the disease occurs every year, usually during the wet season when *Aedes mosquitoes'* populations are high and the rainfall is optimal for breeding. *Aedes aegypti* is closely associated with humans and their dwellings. People not only provide the mosquitoes with blood meals but also water-holding containers in and around the home where the mosquito lays her eggs. In addition, these countries are at periodic risk for epidemic dengue (when large numbers of people become infected during a short period), which require a coincidence of large numbers of vector mosquitoes, large numbers of people with no immunity to one of the serotypes, and the opportunity for contact between the two (CDC 2014).

1.4 Dengue in Brazil

1.4.1 history

The first cases of DF in Brazil during the 20th century were documented based solely on clinical criteria in 1923 in Niterói, Rio de Janeiro (PONTES; RUFFINO-NETO, 1994). New cases were only detected during the period 1981-1982 in the city of Boa Vista, Roraima (in the northern region of Amazonia) with the identification of serotypes DENV-1 and DENV-4 (OSANAI, 1984). Probably this was the expansion of the epidemic wave that hit several countries of Central America and Northern South America in the late 1970s (BRATHWAITE, 2012). The relative geographic and economic isolation of the city apparently helped to limit the epidemic, since there was no viral dissemination to other regions the country. There are also no reports of dengue endemicity in the area after the outbreak.

An epidemic by DENV-1 occurred in 1986 in the metropolitan area of Rio de Janeiro (In the Southeastern region of Brazil). This region was densely populated, with serious urban infrastructure problems, and also concentrated economic activities and human population flow (migratory and touristic) (SCHATZMAYR, 1986; NOGUEIRA, 1990). In the same year, the serotype disseminated to other states of the country, with establishment of autochthonous foci in the states of Alagoas and Ceará. In 1987, DENV-1 reached another four states: Bahia, Minas Gerais, Pernambuco and São Paulo. Then, the serotype arrived in the state of Mato

Grosso do Sur in 1990, hit again with great intensity the State of São Paulo in 1991 and was introduced in the State of Tocantins and Mato Grosso in 1992.

The first autochthonous cases of DENV-2 and DENV-3 were detected respectively in 1990 and 2000 in the metropolitan area of Rio de Janeiro (NOGUEIRA, 2007). The introduction of DENV-2 was associated with the first epidemic of severe dengue in the country. Later in 2008, DENV-4 emerged once again in Brazil and was responsible for several outbreaks during 2010 and 2011 (NUNES, 2012). Remarkably, Bastos et al. (2012) detected in 2011 the simultaneous circulation of all four dengue serotypes in the city of Manaus (located in the middle of the Amazon rain forest in the northern State of Amazonas). Dengue has become endemic-epidemic in most of the states where it has been introduced.

A total of 9,678,709 numbers of cases have been reported in the period 1986-2013. During the period 1986–2006 the country accounts for most of dengue cases in the regions of the Americas, roughly 60% (NOGUEIRA et al. 2007). Teixeira et al. (2013) highlighted that the overall increase in dengue disease was accompanied by a rise in the proportion of severe cases and their trend analysis suggests a worsening of the problem over time. Correspondingly, in 2013, 1,451,432 cases were reported in Brazil corresponding to a 62.5% of the total records of the continent. The same year 6,969 cases were severe and 545 people died of the disease.

1.4.2 Dengue in the State of São Paulo

In 1986, a total of 32 DENV-1 imported cases were documented in the State of São Paulo (28 cases imported from the state of Rio de Janeiro, two from Alagoas and two from Ceará). In 1987, the first dengue outbreaks in the state occurred in the cities of Guararapes (30 reported cases) and Araçatuba (16 reported cases). In the same year 265 cases were imported from the state of Rio de Janeiro, nine from Alagoas and two from Ceará. The following years only imported cases were documented. During 1990-1991 a major DENV-1 epidemic started in Ribeirão Preto and hit several cities in the state. Since 1990 successive dengue epidemics have occurred (PONTES; RUFFINO-NETO, 1994). The detections of autochthonous cases of the other serotypes further aggravated the epidemiological situation: DEN-2 was introduced in 1996 (ROCCO et al. 1998), DENV-3 in 2000 and DENV-4 in

2011.

The state of São Paulo accounted for 13.9% of the total cases (201,498 records) of 2013. The most affected cities were Andradina, Barretos, Bauru, Campinas, Cubatão, Cruzeiro, Leme, Presidente Prudente, Ribeirão Preto, Santos, São Jose de Rio Preto and São Vicente.

1.4.3 Reinfestation in Brazil and in the State of São Paulo

The disease was practically nonexistent by virtue of combating *Aedes aegypti* during the WHO eradication campaign of yellow fever (BRATHWAITE, 2012). The mosquito was absent in Brazil during the periods 1958 – 1966 (FRAIHA, 1968) and 1973 - 1975 (PAHO, 2001).

The first infestation in Brazil by *Aedes aegypti* was recorded in 1967 in the city of Belém (State of Pará) and later in 1968 in the city of St. Louis (State of Maranhão); both foci were eradicated by 1973. But in 1976 the vector was reported in the city of Salvador (State of Bahia), in 1977 in the city of Rio de Janeiro (State of Rio de Janeiro) and finally in several other states. At the same time new foci were found in foreign areas that limit with the country. By 1980, 42 cities were infested and by 1990, the number reached 481. Crucially, at the same time the mosquito *Aedes albopictus* infested 349 cities (this vector was introduced into the country in 1986), and both vectors were thriving in at least 100 cities (PONTES; RUFFINO-NETO, 1994).

In the state of São Paulo, *Aedes aegypti* reinfestation occurred in 1980 in the harbor area of the city of Santos but the foci were eradicated in the same year. In 1985, during the wide spread of the vector in the country, 12 cities (2.1%) of the state were affected. Given that situation, in 1985 the state of São Paulo initiated the "Dengue and Yellow fever vector control program of the State of São Paulo" (Programa de Controle dos Vetores da Dengue e da Febre Amarela no Estado de São Paulo) under the responsibility of the Superintendence of Infectious Disease Control SUCEN (Superintendência de Controle de Endemias). This program monitored the vector infestation indices to estimate epidemic risk and to prevent the spread to non-infested cities but there was not an eradication goal to be achieved. In 1988 some functions were attributed to the cities, especially the mosquito control inside the households. Although control activities were done, there were an increased number of cities

with household infestation. By the end of 1990 the mosquito infested 321 cities in the state (56.1 %) (SECRETARIA DE ESTADO DA SAÚDE DE SÃO PAULO, 1989, 1991a, b, c, d). The geographical progression of infestation occurred from west to east. Critically, in 1986 the mosquito vector *Aedes albopictus* was introduced from the state of Rio de Janeiro into the Paraíba Valley (A geographical region shared for both states), and later appeared in the city of Ribeirão Preto and the metropolitan area of Campinas. This vector moved from east to west. The whole scenario illustrates the failure of the control actions on stopping the spread of the mosquito vector, and therefore of dengue in the State (PONTES; RUFFINO-NETO, 1994).

1.5 Urban dengue outbreaks

There is no specific treatment for DF and no vaccine to protect against the infection. At present, the only method to control or prevent the transmission of dengue virus is to combat vector mosquitoes. In this context, the pattern of spread of DENV in urban areas is of interest because the results may guide the allocation of scarce resources toward vector control.

In addition to the impact on health caused by dengue epidemics, the disease imposes substantial costs on the overall economy of cities: patients are unable to work, children miss school, healthy employees need to stay at home to care for children or relatives, and fear of contagion may keep tourists away (SHEPARD et al., 2011; WETTSTEIN et al., 2012). Programs to control the mosquito population also strain public resources. Suaya et al. (2009) estimated that students lost 5.6 days of school, whereas those working lost 9.9 work days per average dengue episode. Besides, there is a major social impact in those countries where large epidemics occur, often disrupting primary care for hospitalized patients (GUBLER 2012).

Different epidemiological studies done on urban areas in Southeast Asia highlight the importance of geographic and socio-demographic factors in disease transmission, for example, the important relationship between human population density and the rate of dispersal of DENV (densely populated regions operate as major foci of transmission) (CUONG et al., 2013; JEEFOO et al., 2011; RAGHWANI et al., 2011; SCHREIBER et al., 2009; VAZQUEZ-PROKOPEC et al., 2010). Additional ecological factors (seasonality, vector dynamics, and

spatial structure) are also important to improve estimates of diffusion dynamics (RASMUSSEN et al., 2014). Similar studies in Brazilian cities are very limited and most of them are solely based on serological prevalence and incidence data (BARRETO, 2008; TEIXEIRA et al. 2002, 2011). So far, only one work addresses the dynamics of an urban dengue outbreak in the Brazilian city of São José de Rio Preto using viral genetic data (MONDINI et al. 2009). Each city represents a different challenge (depending on different aspects such as infrastructure, economy, dengue awareness and human behavior) for understanding epidemics of dengue. In that sense, exploring diverse urban settings is crucial to elucidate factors associated with the foundation of outbreaks.

2 OBJECTIVES

The overall research goal is to elucidate the pattern of urban spread of DENV at the city of Guarujá, São Paulo. The specific objectives are:

- (i) To identify the serotypes and genotypes of viral strains;
- (ii) To infer the evolutionary history of the circulating strains;
- (iii) To infer the history of phylogeographic distribution using established methods;
- (iv) To identify demographic factors and / or socioeconomic associated with the onset of the outbreak;

3 METHODS

Guarujá (23°59'37"S 46°15'23"W) (Figure 5) is a municipality from the State of São Paulo. It is a coastal city embedded in the Serra do Mar coastal tropical rain forest in Santo Amaro Island whose main economic sources are seasonal tourism and port related activities. The estimated population in 2013 was 306,683, the population density was 2,026.80 inh/km² and the area is 143 km². The city has a tropical humid climate ("Af" classification), which is characterized by having higher average air temperature and rainfall. The average annual temperature is 24.7 °C (Min 18 °C | Max 31.3 °C) and the annual rainfall is 3,413 mm (Min 155.6 mm | Max 412.8 mm); February is the wettest month (rainfall of 412.8 mm) and August is the less rainy one (155.6 mm) (CENTRO DE PESQUISAS METEOROLÓGICAS E CLIMÁTICAS APLICADAS À AGRICULTURA, 2014). Since 1997 the Epidemiological Surveillance Center Of Guarujá have been reporting autochthonous dengue cases. In 1999, an DENV-1 epidemic hit several regions of the state of São Paulo, approximately hundred cities, among which stood out Santos, São José do Rio Preto, São Vicente, Cubatão and Guarujá, that were responsible for over 70% of cases statewide.

3.1 Sample collection

Sample collection was done with patients seen in Santo Amaro Hospital (HSA), the Basic Health Units (Unidades Básicas de Saúde, UBS), the Emergency Care Units (Unidades de Pronto Atendimento, UPA) and the Family Health Units (Unidades de Saúde da Família, USAFA) from Guarujá, according to the established guidelines by the Brazilian Ministry of Health.

Patients of any age with symptoms and signs of dengue disease were included. Symptoms included fever, frontal or retro-orbital headache, severe bony and myalgic pain in legs, joints and lower back, abdominal pain, nausea, vomiting, taste disturbance and anorexia. Signs included high fever (usually between 38.5 and 41°C) persisting for 24 hours, rash, hemorrhagic manifestations, hypotension and narrow pulse pressure.

Figure 5 - Territorial division of Guarujá.

Guarujá is a municipality in the São Paulo state of Brazil. The population is highly urbanized. The urban area of Guarujá is divided in four administrative regions and 29 neighborhoods. Much of the region on the coast (mainly Asturias Pitangueiras, Enseada, and Tombo) has buildings devoted to the seasonal population (summer). Tourism, therefore, is the most important component that moves the economy. The other part, relevant and non-seasonal, comes from the port and related activities, such as transportation.

After the patient met the requirements and signed a consent form, a blood sample was collected and used in a Dengue Duo Test. This test is a qualitative immunoassay for the simultaneous detection of the NS1 antigen as both the IgG and IgM antibodies. During the acute phase of dengue infection, the presence of IgM antibodies alone suggests primary infection (HALSTEAD, 2007). Dengue NS1 antigen, as reviewed by Kassim et al. (2011), is a highly conserved glycoprotein produced in both membrane-associated and secretion forms. It is abundant in the serum of patients during the early stages of DENV infection and can be detected before the formation of antibodies. Therefore, to increase the chances of processing positive samples, only the serums of patients with positive IgM and/or NS1 outcomes were referred to our laboratory.

3.2 Molecular testing

The viral RNA was extracted from the serum samples with the QIAmp viral RNA mini kit (Qiagen, Venlo, Limburg, Netherlands) according to the manufacturer's instructions. A first RT-PCR was performed using Dengue generic primers for the capsid/premembrane junction (C-preM), which is a conserved region in dengue viruses and used in reference laboratories for detection. This diagnostic RT-PCR was done using extracted RNA or newly synthesized cDNA using the SuperScript® VILO™ cDNA Synthesis Kit (Life Technologies, Carlsbad, California, United States). For serotype typing a second nested PCR (See Figure 6) or DNA sequencing was used. In any of the described amplification procedures any of the following polymerase enzymes were used: GoTaq® Green Master Mix (Promega, Madison, Wisconsin, United States) or SuperScript™ III One-Step RT-PCR System with Platinum® Taq (Life Technologies).

Figure 6 - Typing of Dengue viruses.

Amplified cDNA obtained from the reverse transcription is used as a template for sequencing or in a second PCR reaction (i.e. nested PCR) to determine the Dengue serotype. EHD1 and all four serotype-specific primers are added to this second PCR reaction mixture. However, only one electrophoretic band should be observed, its size depending on the serotype present in the sample

The amplification of the envelope gene of positive samples was done using previously published primers. The ExoSAP-IT reagent was used for PCR Product Cleanup (Affymetrix, Santa Clara, California, United States). The sequencing reaction was performed using BigDye® Terminator v3.1 Cycle Sequencing Kit (Life Technologies) and the sequencing product was purified using the BigDye XTerminator Purification Kit (Life Technologies) and sequenced on an ABI PRISM® 3130 Genetic Analyzer (Life Technologies). Contigs were assembled using the program Codon Code aligner.

3.3 Genetic analysis

Polymorphisms were analyzed with the R-package APE and the program DNASP v5. A Tajima D statistic test, as proposed by Tajima was used to evaluate deviations from the neutral expectation of molecular evolution.

We screened for recombination using an array of tools available in the HyPhy Package 2.2 (SBP-Single Breakpoint Recombination and GARD-Genetic Algorithms for Recombination Detection), SplitsTree 4.12.6 (pairwise homoplasy index, PHI) and RDP-Recombination Detection Program version 4.36 (RDP, MaxChi, Chimaera, Geneconv, Bootscan and Siscan).

The haplotype dataset was used in a selection analysis to determine whether amino acid positions were subject to negative or positive selection pressures. Analyses were done using Guarujá samples only or using an outgroup (dengue strain H7880090 isolated in Boa Vista, RR-Brazil). This strain belongs to genotype II and represents one of the DENV-4 original sequenced isolates introduced into the country (29 of Nov 2010). Both genealogy-based, codon-site models Single Likelihood Ancestor Counting (SLAC) and the fixed effects likelihood (FEL) available in the HyPhy package (www.datamonkey.org) were used to estimate the non-synonymous (dN) and synonymous (dS) rates of substitution. We used an inferred ML tree with GARLI as an input (POON et al., 2009).

Counting methods (*e.g.*, single likelihood ancestor counting, SLAC) are the most efficient and are well suited to analyzing large data sets (*i.e.* over 40 sequences), but can be more conservative than other methods. On the other hand, FEL methods, proposed by Suzuki and Gojobori (1999), estimate site-specific rates directly from each site independently using maximum likelihood methods, and are far more time-consuming but more sensitive, and may be more successful at detecting selection in smaller data sets.

3.4 Phylogenetic reconstruction

Sequences were aligned using Muscle software (EDGAR 2004a, b) followed by visual inspection and manual editing with the BioEdit v7.2.5 (HALL, 1999) or Mesquite 2.75 (MADISON; MADISON 2014).

The JModeltest software was used for the statistical selection of models of nucleotide substitution under Akaike information criterion (AIC) (DARRIBA et al. 2012; GUINDON; GASCUEL, 2003).

The dataset generated was used for phylogenetic reconstruction using both the maximum likelihood criterion (ML) and Bayesian Inference (IB). The program GARLI version 0.95 (ZWICKL, 2006) was used for the ML analysis. The genetic algorithm method implemented in GARLI estimates simultaneously the best topology, branch lengths and the best values for the model parameters. Bootstrap analyses with 100 replicates were done to place confidence values on grouping within the trees.

Bayesian maximum clade credibility trees were inferred from a set of plausible trees sampled at the stationary phase of eight independent Markov Chain Monte Carlo (MCMC) runs with 200 million generations each using the estimated model and a strict or relaxed (uncorrelated lognormal) molecular clock (DRUMMOND et al., 2006). The values of the AIC were used as the initial values for the substitution model in the Bayesian analyses but estimated together with a proportion of invariants parameter. Analyses were done with Beast v2.1.3 (BOUCKAERT et al., 2014). Convergence of parameters was assessed using Tracer v1.6 program (<http://tree.bio.ed.ac.uk/software/tracer/>) until all parameters estimates showed Effective Sample Size (ESS) values over 200.

In a previous study we studied the Evolutionary history of Dengue virus following a Bayesian phylogenetic approach using 310 isolates sampled from 37 countries during the time period 1956-2008 (VILLABONA-ARENAS; ZANOTTO, 2011). The substitution rate estimated for that data using a relaxed molecular clock model was 9.1×10^{-4} subs/site/year (HPD 7.6×10^{-4} to 1.1×10^{-3}). Because we have prior information on the overall dengue evolutionary rate, we reflected that in our estimation with a prior distribution on the ucl.mean that is the rate parameter of the model (ucl.mean is the simple arithmetic mean of the branch rates: If r_i is the rate on the i 'th branch then ucl.mean is simply the sum of r_i for all i divided by the number of branches ($2n-2$)).

Beast does not allow specifying a minimum span time for two sequences to be related; hence in some cases the Time to Most Recent Common Ancestor (TMRCA) can be estimated at a value that is less than the value that is biologically plausible. This is the case of a dengue transmission chain where we have hosts and vectors (a vector and a human incubation period) but we only sampled from the hosts. The estimation is only problematic

when dealing with dengue sequences sampled on a daily basis: given the low diversity, several internal nodes are inferred to be closer in time and this may not reflect accurate transmission cycles. To try to account for this and check if the time unit is influencing the parameter estimation, we proposed different schemes based of the DENV incubation periods. The best-fitting model for dengue transmission estimated that 95% of mosquito incubation periods are between two and 15 days, with mean of 6.5 days; the human estimates are 5.9 days, with 95% expected between days three and 10 (Figure 3). Therefore, once a person is infected, it takes some time before he can transmit the virus to a mosquito taking a blood meal, and also takes some time before the virus crosses the gut epithelium and reaches salivary glands for transmission. Given so, our time schemes were: (i) consider the whole sequences to be contemporaneous and, (ii) the average time (12 days, $5.9 + 6.5$), (iii) the shortest time (5 days) and (iv) the longest time (25 days) for a successful transmission event to occur.

3.5 Phylodynamics

Birth–death skyline (BDSKY) and SIR (BDSIR) models were used as population priors for the phylodynamics approach (KÜHNERT et al., 2014; STADLER et al., 2013). We did not use standard coalescent models because they require the population size to be constant or to vary deterministically and also requires that the effective size of the population be large enough for the probability of two coalescent events in a single generation to be very small. To accommodate stochastic population size changes within the phylogenetic reconstruction, a tree prior based on a birth–death process is more appropriate. The BDSKY allows for serially sampled data and rate changes over time. The main parameters to consider under this model are the basic reproduction number, the ‘become uninfected’ proportion, the birth and death rates, the sampling proportion, and the origins parameters.

The basic reproduction number, R_0 , is the number of secondary infections that results from a single infected human. The infection will die out in the long run if this value is below 1.0. For this parameter we use a prior distribution $LogN(1,1)$. Dengue Infected people are infectious during approximately a five days period when large amounts of virus are in the blood, the ‘become uninfected’ parameter, γ , is the inverse of the infectious period. We used a prior that assumes an infectious period of roughly 5 days: $LogN(-4,1)$. In the birth-

death models, the sampling proportion, s , is directly modeled as a stochastic process. Most dengue infections are unapparent or result in a self-limited febrile illness (BURKE 1988). Moreover, in areas where multiple serotypes of dengue are endemic and infection rates among children are sufficiently high to immunize the population before adulthood, only a small proportion of infections result in disease. Given so, even when we have a densely sampled outbreak, we assume the proportion of observed/sampled infections is rather small when compared to the overall number of infections at each sampling time. Since Birth, Death and Sampling are correlated parameters it is helpful to put some prior information on the later. Given so, for this parameter, we use a prior distribution $Beta(2,10)$. The origin parameter represents the onset of the epidemics and it is larger than the tree height. Given so, for the origin parameters we used a uniform distribution that ranges from 160 up to 365 days.

The BDSKY can be parameterized to enable the underlying population dynamics to be modeled as a compartmental susceptible–infected–removed (SIR) model, which accounts for changing host population composition. In this model we calculate the parameters dS and dR that refer to the changes in S (susceptible) and (removed) over time. The model starts with one infected ($I = 1$) and then get the changes in I over time by computing $dI = dS - dR$. Prevalence is equivalent to I , while incidence is given by dI . In BDSKY we calculated piecewise constant rates that were used in the BDSIR in order to incorporate host population dynamics. For the initial susceptible population size parameter $nS(0)$, we proposed a prior distribution with a mean around Guarujá total population $LogN(7,1)$ as shown in the study site section.

The calculation of the epidemiological parameters for the dengue outbreak, the rate of evolutionary change (μ) (subs/site/year) and TMRCA were estimated during the Bayesian inference. Sequences were dated according to the date of sampling. The confidence intervals for each parameter were given by the 95% highest probability density (HPD). Because the priors were not necessarily the same, different models were compared calculating Bayes factor in the program Tracer v1.6.

3.6 Spatio-temporal dispersion pattern

The Beast program allows for an ancestral reconstruction of continuous/discrete states where the spatial diffusion of the time scaled genealogy is modeled as a continuous-time Markov chain (CTMC) process over continuous/discrete sampling locations (LEMEY et al., 2009, 2010). We used the Bayesian stochastic search variable selection (BSSVS) approach, which assumes exchange rates in the CTMC to be zero with some prior probability, to find the most parsimonious set of rates explaining the diffusion process along the phylogenies for the geo-referenced dataset. We use the software Spread v1.0.4 to visualize diffusion rates over time: locations were represented by neighborhoods (discrete phylogeography) or the latitude and longitude of the household addresses (continuous phylogeography) (BIELEJEC et al., 2011). A Bayes factor (BF) test was run to identify the rates contributing to the migration path.

3.7 Transmission Tree

We used the R-package *outbreaker* (JOMBART et al., 2014) to reconstruct the transmission tree and infer the number of imported cases. This package implements a Bayesian approach for reconstructing outbreak data from densely sampled pathogen sequences. We estimated the evolutionary rates and epidemiological parameters (incidence, R_0 and s).

For the estimation of the transmission tree we specified a generation time distribution (w.dens) and a colonization time distribution (f.dens) that was in agreement with daily sampled vector borne outbreaks. If the generation time is the time that has elapsed between one person being infected and that person infecting someone else, then for dengue this is equivalent to the sum of the length of both incubation periods. In this case, we proposed a normal distribution with a mean around 12.4 days and a range between five and 25 days, $N(12,1)$ (Figure 3). For the f.dens the rationale was as follows: the time during which the virus can be sampled from the patient and the level of viremia usually follow each other closely. Viremia is typically highest in the first three to four days after

onset of fever but then falls quickly to undetectable levels over the next few days. Hence, we proposed an $f.dens\ N(4,1)$.

3.8 Extrinsic Traits

For more detailed information on demographic and/or socioeconomic factors associated with the outbreak it was inferred the number of basic reproduction (R_0) for each patient that approximates the value of "contagion" of individuals. These values were correlated with factors such as age, gender and variables obtained from residents of neighborhoods.

4 RESULTS AND DISCUSSION

4.1 Sampling

1216 serum samples were collected in the city of Guarujá for a period of 160 days, from January 2nd to June 10th of 2013 (Figure 7). 338 sequences were not shipped or not processed because of technical problems (*e.g.* very low volume of sample). 879 samples were tested in our laboratory for dengue: 525 (60% of processed samples) were positive for dengue and 354 were negative.

We were able to geo-locate 286 patients (81%) based on the addresses provided, from the cohort of 354 patients (Figure 8). The reasons for not geo-coding the other 68 patients were the lack of their complete home address in the records from the Guarujá Municipal Health Department.

Figure 7 - Dengue sampling during 2013

Serum samples were collected in a daily basis in different health units of Guarujá. Only samples with positive IgM and/or NS1 Dengue Duo Test outcomes were referred to our laboratory. "A" shows the total sampling and "B" the geo-referenced sampling.

Figure 8 - Geo-referenced samples

Each mark represents a sample and the household of the patient. The colors represent the relative time of collection following the spectrum of visible color (e.g. red color represents the oldest samples and the blue color represents the most recent ones).

4.2 Hyperendemicity of dengue

Positive samples from the onset of the epidemic were used for phylogenetic reconstruction to identify the circulating serotype. The capsid/premembrane junction that was proposed by Lanciotti et al. (1992) et al. for typing DENV was amplified and sequenced.

A small time-stamped dataset comprising 35 sequences that were representative of both serotypes and genotypes was retrieved from GenBank and aligned with our sequences. A phylogenetic tree (Figure 9) was built using a Bayesian approach, and our evolutionary estimates matched those of Twiddy, Holmes and Rambaut (2003), which validated our analysis.

Figure 9 - Phylogenetic relationships among dengue serotypes and strains isolated during the beginning of the outbreak in 2013.

The Maximum Likelihood tree was inferred using a 511 bp region from the C-prM junction. Samples from Guarujá belonged to the four serotypes. The figure is presented as a cladogram, for illustrative purposes.

505 (96% of positive samples) were positive for DENV-4 and 20 belong to any of the other three serotypes (10 DENV-1, 8 DENV-2 and 2 DENV-3).

The presence of the four serotypes in a single outbreak in one of the most densely populated areas of Brazil is a disturbing finding that has also been documented in Asian countries, particularly in India, which has the largest dengue burden in the world. Recurring dengue epidemics in that country resulted in the establishment of hyperendemic areas,

typically in large, densely populated cities, where most DENV serotypes circulate in a sustained fashion (CHAKRAVARTI; ARORA; LUXEMBURGER, 2012). Roughly two years after the report on the presence of DENV-4 in Manaus, we now find the four serotypes co-circulating in the south of the country in a municipality of São Paulo. Given so, we expect a gradual change in the dengue epidemiological pattern because hyperendemicity and high infection rates lead to immune populations before adulthood. Hence, the new generations of children become the ones at risk during the next outbreaks (INNIS, 1995). For example, prior to 2007, the disease affected mostly adults (20- to 40- year-old people) but during the 2007–2008 epidemics, over 53% of the cases affected children under 15 years of age (TEIXEIRA et al., 2008). Consistently, Rodriguez-Barraquer et al. (2011) argued that the disease's shift towards younger patients observed in Brazil could be partially explained by the accumulation of immunity against multiple serotypes in older people.

We were interested in the phylogeographic of DENV during an urban outbreak so we proceed to obtain the envelope gene sequences for DENV-4 (1485 bp-long) that was the most prevalent serotype (Figure 10). We obtained 354 (70% of positive samples) sequences and proceed with the computational analyses with this dataset.

Figure 10 - Amplification of the DENV-4 envelope gene.

The envelope gene was amplified using two sets of primers. The same sets of primes were used for sequencing. Additional primers for sequencing were used, as described by FOSTER et al., (2003), when the quality of the reads was not good during contig assembling.

4.3 DENV-4 sequence analysis

The 7.3% of the envelope gene have variable (polymorphic) sites, 4.4% being singletons variable sites and 2.9% parsimony informative sites. The Tajima's D value for DENV-4 was of -2.56 ($P < 0.001$) and signifies an excess of low frequency polymorphisms relative to expectation, indicating population size expansion and/or purifying selection (most likely due to functional constraints imposed on the viral protein). Studies involving the estimation of dN/dS ratios on the envelope gene revealed extensive purifying selection (HOLMES, 2003, KLUNGTHONG et al., 2004, ZHANG et al., 2006 and COSTA; VOLOCH; SCHRAGO, 2012). In agreement with this, the overall dN/dS value for the entire DENV-4 polyprotein-coding region was 0.15 (estimated 95% CI from 0.10 to 0.21) that implied purifying (stabilizing) selection. There was no evidence for adaptive evolution in any codon. Figure 11 shows the sites that experienced significant purifying selection and they all fall in the central and dimerization envelope gene domain.

Figure 11 - Selection analysis of the DENV-4 sequences from Guarujá

The “A” plot illustrates the single nucleotide polymorphisms over the entire envelope gene and the conserved domains of the Flavivirus glycoprotein. “B” and “C” plots illustrate the negatively selected sites as detected respectively by the SLAC and FEL approaches.

4.4 Evolutionary relationships of DENV-4

The Tamura-Nei model (TrN, variable base frequencies, equal transversion rates, variable transition rates) of nucleotide evolution with Gamma-distributed variable rates (Γ) was the best-fit model for the data.

A total of 354 sequences of DENV-4 strains isolated in the study and reference sequences from diverse geographical locations were utilized for the construction of a ML tree (Figure 12). The phylogenetic tree evidenced that All DENV-4 strains isolated during this

study belong to genotype IIA. As reviewed by Yamashita et al. (2013) this genotype was initially circulating in Southeast Asia and several Pacific countries during the 1970s. The first cases were reported in Indonesia in 1976–1977, Tahiti in 1979, and New Caledonia in 1981. Then, in 1980s, the virus emerged in the Neotropics, with first reports coming from Colombia, Brazil and French Polynesia. Then, many cases with this genotype were reported from other countries around the region, including the USA, Venezuela, Ecuador, and finally stroke once again Brazil.

Figure 12 - Phylogenetic relationships among representative samples of different genotypes and local isolates of DENV-4 during the dengue outbreak in the city of Guarujá.

Maximum Likelihood tree inferred using haplotype data and the envelope gene. Names have been removed and the relationships presented as cladogram for clarity. All the isolated belonged to genotype 2A.

Figure 13 shows the summarized MCC tree estimated with a BDSKY model. Two major lineages (A and B) were identified. The co-circulation of multiples lineages is a frequent finding during dengue outbreaks (MONDINI et al., 2009; RAGHWANI et al., 2011). That these lineages emerged at approximately the same time and dispersed over similar spatial regions suggests that they are of broadly equivalent fitness. SCHREIBER et al., (2009)

using genomic data from an outbreak in Singapore showed that two serotypes spread concurrently and with similar epidemiological and evolutionary profiles. In that study also, there was no evidence for adaptive evolution. These findings indicated that ecological or immunological rather than virological factors were the key determinants of the observed epidemic dynamics.

There were not major topological or parameter estimation differences among the MCC tree in Figure 13 and those MCC trees inferred using the average time and shortest time for EIP + IIP time schemes. When treating all the sequences as contemporaneous or using the longest time for EIP + IIP as a time scheme, there were problems with convergence during the runs, the estimated average rate of evolutionary change tended to be very high (*e.g.*, 1.02×10^{-3} nucleotide substitutions per site, per day) These findings pointed out that given the low level of diversity, the temporal information is needed to achieve misleading inferences.

The mean evolutionary rate estimated with BEAST approach was 7×10^{-6} (95% HPD interval 4.6×10^{-6} , 9.4×10^{-6}) substitutions per site per day. The rate was similar to that obtained with a strict molecular clock 7.1×10^{-6} (95% HPD interval 4.6×10^{-6} , 9.7×10^{-6}).

Figure 13 - MCC tree estimated with a BDSKY model.

Maximum clade credibility tree estimated using the envelope gene, a relaxed clock model and a Birth-Death epidemiological model. Name colors represent isolation dates. Two major lineages are indicated as “A” and “B”.

4.5 Epidemiological findings

Figure 14, 15 and 16 show the epidemiological results for both, the BDSKY and the BDSIR models.

A $R_0 > 1$ points out that the infection will be able to spread in a population following exponential growth. The R_0 estimate with the BDskyline model had a mean value of 1.74. The R_0 estimate remained above 1.0 during most of the outbreak, decreased below 1.0 only by the end of March, and increased approximately to 2.0 and 3.0 during January and February at the thick of the summer season. Basic reproduction numbers estimated for dengue range between 1.33 and 11.6 (HALSTEAD, 2008). Favier et al. (2006) and Massad et al. (2003) analyzed epidemic data from several cities in Brazil and developed values ranging from 3.8 to 5.1 and from 2.7 to 11.6, respectively. A More recent study in Brazil reported values between 1.5 and 2.75 (PINHO et al., 2010) that agree with our current estimate. The estimate with the BDSIR model had a mean value around 1.6 and the reconstructed dynamic over time (Figure 15) pointed out that the epidemic was over in April. Nevertheless, I will

later show in the text, that the tail of the distribution of R_0 shows that a small of individuals are responsible for a large number of secondary transmission events. This “epidemiological core” is therefore relevant for the overall dynamics of the epidemics.

Low herd immunity thresholds characterize diseases with high R_0 numbers (A herd immunity threshold is the maximum level of immunity beyond which transmission is eliminated). It is crucial to consider that the population of Guarujá is a *naïve immune population* to DENV-4 infection but repeatedly exposed to the other serotypes, hence theoretically cross-reaction immunity from previous exposures may contribute to prevent higher R_0 values.

The estimated sampling proportion (Figure 14) changed throughout the outbreak and with the highest value around March and specially low during the month of April. As showed in Figure 7, April represents a gap in the sampling and we will retake this later in the discussion.

Figure 14 - DENV-4-1 in Guarujá 2013: Birth–death skyline plot

Median estimates and HPD intervals for the (upper) effective reproductive number and (lower) sampling proportions. Inferences were based in a birth-death epidemiological model and a relaxed clock model.

Figure 15 - Coarse-grained reconstructed effective reproduction number using BDSIR models.

The median effective reproduction ratio, computed from the posterior birth–death rates and SIR trajectories, shows the epidemic fading out since March. An initial value of $R_0 = 1.8$ was estimated during the simulation. During this reconstruction the time when R_0 reached values of 1 or lower was determined to be at 90 days (*i.e.*, end of March) after the outset of the epidemics.

The SIR model provides trajectories (Figure 16) of the number of susceptible, infected and removed individuals. Based on the trajectory of infected individuals, the average transmission rate in short time intervals throughout the epidemic was determined. Even when the infected trajectory conforms to the epidemiological situation observed in Guarujá, the removed and susceptible trajectories seem to be at the order of magnitude of the actual demographic size and showed that the epidemic started to decline in March. I will argue that these figures may be realistic based on the following epidemiological evidence. For example, in the city of Salvador, Brazil the observed seroincidence in areas with low premise index was around 54.6%, which led the authors to suggest the existence of a silent epidemic during a period perceived by the Health Services as of low endemicity, pointing out to the strength and speed of dengue transmission in the city (TEIXEIRA et al., 2002). Moreover, two seroepidemiological studies in Singapore estimated an innapparent dengue rate of 95.7% (1 out of 23 cases were diagnosed and notified) and 78% (YAP et al 2013; YEW et al. 2009,). Finally, values as high as 99.78% were obtained from our seroprevalence study in the nearby city of Jundiaí in 2010 (VILLABONA-ARENAS et al., 2012). Therefore, the results presented here are consistent with the findings of previous serological surveys as well as findings in other countries that a majority of dengue virus infection is either asymptomatic

or subclinical, and notified cases only represent a fraction of the actual extent of infection in the population (BECKETT et al., 2005; BURKE et al., 1988; CHEN et al., 1996). Based on previous estimates, the higher values in our trajectories (Figure 16) may well be attributed to unreported cases, including those that are inapparent. A total of 1216 suspected dengue samples were collected during the current study and, using a percentage as low as 4.3% for the reported clinical cases, we calculated that the actual number of infections would be around 28000. Nevertheless, further improved model parameterization is needed to check for convergence of results and critically, a seroepidemiology of dengue virus infection in a representative sample of the resident population will be necessary to confirm any estimate of the proportion of asymptomatic dengue infection.

Figure 16 - SIR trajectories and incidence of Dengue from the Guarujá Outbreak

Bayesian posterior median trajectories: the overall SIR dynamics (upper plot) show that the epidemic peaked around March and declined since. Zooming into the number of infected, i.e. the prevalence over time in (lower) enables comparison to the incidence.

4.6 Spread dynamics

Figure 17-19 show the main finding with the phylogeographic analyses. The ancestral state reconstruction in the Location-annotated MCC tree from Figure 17 shows that early in the outbreak the virus was already spreading to many locations, this is consistent with our epidemic temporal outset estimate (origin), the initial incubation periods of dengue virus and the high innapparent dengue proportions. The start of the 2013 epidemic was

pinpointed at the last two weeks of November 2012 (205 days back in the past 95%HPD interval 198, 212.5), precisely when the municipality led the ranking of infestation of dengue in the State of São Paulo.

Figure 17 - Location-annotated MCC tree

Maximum clade credibility tree inferred using the envelope gene, a strict clock model and a birth-death epidemiological model. The concentric ring depicts the temporal component of the reconstruction.

During the onset of the epidemic, in which there is mainly silent transmissions of dengue, the virus is spreading to all locations where the vector mosquito was thriving. The first record of a clinical dengue case with a Dengue Duo test was on December the 20th 2012 from the Cachoeira neighborhood and seven more cases were documented in the next days before 2013 in Pae Cará, Morrinhos, Vila Zilda e Mar e Céu. Our first geo-referenced sample was collected in January the 2nd 2013, at the Enseada neighborhood. Following the phylogeographic approach, this place appears as one of the major transmission foci to neighboring areas and to Pae Cará. Enseada and Pae Cará, important hubs of Denv-4 acitvity,

have in common low income and high population densities. Following suit, Pae Cará become also an important source of virus to neighboring areas (Figure 18). Moreover transmission routes become interconnected late in the outbreak (Figure 19), densely crisscrossing the main area of socio-economic integration of the municipality, where administration, popular commerce and mass transportation are centralized.

Figure 18 - DENV-4 Introduction routes to each neighborhood.

Spatial reconstruction done using the location-annotated MCC tree from Figure 17. The thickness of the line indicates the chronological order of events during the onset of the outbreak. These events occurred in a very small time scale (few days). The position of the lines does not indicate a specific route of diffusion, only the connection between the two locations.

Figure 19 also highlights the overall spread dynamics of the outbreak as a whole. With a north-south diffusion processes with Pae-Cará as a major transmission foci, a west-

east diffusion process in the littoral area (through Enseada) and a third one from the administrative region 4 to the neighborhoods in the administrative region number 1. This cardinal pattern fits well with access roads of the island.

Figure 19 - Rates that that explains the DENV-4 diffusion process among neighborhoods.

Spatial reconstruction done following a location-annotated MCC tree from Figure 17. Only rates with a bayes factor value above 3.0 are presented. The position of the lines does not indicate a specific route of diffusion, only the connection between the two locations.

4.7 Transmission dynamics

With the aid of the package outbreak we inferred a transmission network for the outbreak in Guarujá. Figure 20 and figure 21 show the reconstructed transmission tree with and without a time axis.

Figure 20 - Reconstructed transmission tree with time axis

Ancestries with support >50%. Colors represent mean infection dates (red color represents the oldest samples and the blue color represents the most recent ones) and node size reflects individual reproduction numbers.

Figure 21 - Reconstructed transmission tree

Ancestries with support >50%. Colors represent mean infection dates (red color represents the oldest samples and the blue color represents the most recent ones) and node size reflects individual reproduction numbers. Disconnected nodes and networks represent different introductions.

Two lineages were identified in the phylogenetic analysis and are also present here but, different introductions were also documented. Looking for nodes for which the ancestor is inferred as unknown leads to the systematic assessment of imported cases. Following this approach, 10 independent introductions were unveiled and they are shown in Figure 21 as disconnected nodes or small separated networks. The latter represent novel introductions that also reached some level of transmission.

The mean degree distribution for the network was 1.92. This value is in some way equivalent to R_0 estimates, once it tells us about the number of connections that a random selected node has to other nodes. Crucially, this analysis also allowed us to compute the number of secondary cases per infected individual, that is, the individual effective

reproduction numbers (R_i). The node size in the network reflects these reproduction numbers.

The estimated evolutionary rate was 3.61×10^{-5} (Min 2.15×10^{-5} , Max 6.29×10^{-5}) and is higher than the one estimated by BEAST (the highest average rate of evolutionary change recorded for this serotype is 2.94×10^{-6} nucleotide substitutions per site, per day (KLUNGTHONG et al., 2004)). Nonetheless, this approach relies in a very simple nucleotide substitution model (one transition and one transversion rate) and do not take into account rate variation across lineages. Moreover, the chronologies of events during the outbreak were consistent with inferences based on the posterior distribution of rate values.

Figure 22 - Network degree distribution for the reconstructed transmission tree

The degree of a node in a network is the number of connections it has to other nodes and the degree distribution is the probability distribution of these degrees over the whole network. The “0” values point out to different introductions.

In figure 23, the network is plotted over Guarujá with Enseada concentrating many cases with higher R_i . A few important features became evident when looking at Figure 23. Firstly, the main hubs of transmission overlapping with higher population density and connected via the road system become evident as in Figures 18 and 19. Nevertheless a salient point of the geographical depiction of the transmission network was that the adjacent Jardim Acapulco and Pernambuco Beach, high-income, low-population density neighborhoods at the North of the Island had no cases sampled. This is not simply explained by their distance from the main areas of the outbreak near downtown, because

both low-income areas nearby: Favela da Maré (to the South) and Pereque Beach (to the North) had several cases were reported during the entire outbreak. I would argue that the fact that these two high-income neighborhoods had privately-owned and hired vector control services may explain this discontinuity on viral spread.

Figure 23 - Reconstructed transmission tree plotted over Guarujá.

Ancestries with support >50%. Colors represent mean infection dates (red color represents the oldest samples and the blue color represents the most recent ones) and node size reflects individual reproduction numbers. Vertices have been removed for clarity.

4.7.1 Patient attributes and transmission.

R_i is a proxy for the "infectiousness" of individuals; given these values we correlate it to other measured covariates. We did correlate them with factors such as age, sex, Dengue Duo test results, location, and available neighborhood features (population size, number of low income households, average income per household, number of shanty towns, garbage collection and water supply).

For the variable sex we used a t-test and to determine if the two sets of data are significantly different from each other. Briefly, using the Mann-Whitney-Wilcoxon Test, we can decide whether the population distributions are identical without assuming them to follow the normal distribution. For the Dengue Duo test results and locations we used The Kruskal–Wallis one-way analysis of variance, a non-parametric method for testing whether samples originate from the same distribution. This test is used for comparing two or more samples that are independent, and that may have different sample sizes, and extends the Mann–Whitney U test to more than two groups. We also made the correlation with the parametric equivalent of the Kruskal-Wallis test is the one-way analysis of variance (ANOVA). The latter compare means of two or more samples (using the F distribution) with numerical data. For the variable age and the neighborhood features we used the Pearson and Spearman correlations. On the other hand, Spearman is computed on ranks and so depicts monotonic relationships while Pearson is computed on true values and depicts linear relationships. If the Spearman value is higher than the Pearson value, it means that a correlation is monotonic but not linear. Only the correlations with the Dengue Duo results were statistically significant (Figure 24). As expected, there was no clear-cut correlation of number of secondary cases and age of the patient. But the highest concentration of points and highest values were observed at people at school and working ages, which tend to be more mobile and also constitute a high proportion of the active population. Another interesting observation was that infectivity, as indicated by number of secondary cases linked to each patient, was not different among sexes, but males had a significantly higher number of outliers with higher number of secondary cases, compared to women ($c_2 = 9.2$, $df = 1$, $p = 0.0024$). Moreover, we observed a significantly higher number of outliers with high number of secondary cases of individuals that were highly reactive to NS1 and had low reactivity against IgG and IgM ($c_2 = 525.1$, $df = 5$, $p > 0.0001$). We checked if these outliers

could be linked to specific viral lineages, but found them to be completely scattered on the viral genealogy. Therefore, these results could suggest that at the epidemiological level, lower levels of immune response could facilitate transmission. Since immune reaction is related to pathogenesis, this last observation (high NS1 production and low immune reactivity) could also explain to some extent cryptic transmission among asymptomatic carriers.

Different studies investigated relationship between the disease and diverse variables without finding statistical associations that explain the risk of dengue (BARCELLOS et al., 2005; MONDINI; CHIARAVALLOTI NETO, 2007). In contrast, Medronho (1995) associated the disease to slum areas; Chiaravalloti, Moraes and Fernandes (1998) described the increased risk of dengue in areas of higher population density and lower income; Costa and Natal (1998) assume the positive relationship between incidence and population density, garbage collection and sewage, but do not find a statistically significant relationship between dengue and water supply; and Vasconcelos et al., (1999) established the risk in areas of higher income. Machado, Oliveira and Souza-Santos (2009) described the highest risk of dengue in heterogeneous areas located nearby access roads. In Guarujá, we observed that the viral diffusion follows the access roads. These observations indicates the importance of population dynamics in the formation of a large conglomerate risk for dengue

Figure 24 - Inferred infectivity.

The R_i is a proxy for the infectiousness of individuals that can be correlated with other measured covariates.

4.8 Filling the gap

As showed in previous sections there was a gap around April where fewer samples were collected. The official records evidenced that the number of dengue notifications peaked in April but decreased thereafter. Figure 27 shows that the number o cases were growing steadily until the mid or March when community efforts (mutirões, shown by yellow triangles) to remove mosquito-breeding sites in shantytowns or neighborhoods started and

slightly flattened the rise. Nonetheless, in April the number rise quickly but the sampling was almost nonexistent.

Figure 27 - Official reported autochthonous dengue cases reported in Guarujá, 2013.

epidemiological, all patients who experience acute febrile illness lasting up to seven day, accompanied by two or more symptoms, are considered confirmed cases. This not only impacted our sample collection that relied in serological testing, but also diseases with a variety of non-specific signs and symptoms can be notified as dengue further increasing the number of notifications. Moreover, the Coordinator of Epidemiological Surveillance of Guarujá also informed (personal communication) that there were bureaucratic circumstances that limited the sampling for the research. Due to the increased number of cases, it was necessary to direct the sampling to the Surveillance and Control Program of Instituto Adolfo Lutz. This institute contributes to the epidemiological health surveillance in the state.

The coordinator for Dengue Prevention and Control from Guarujá informed (personal communication) that, following the surveillance done with the Dengue Duo test, mosquito control actions were planned and directed to those neighborhoods that were having more incidence during the outbreak (Enseada, Paé Cara and Morrinhos). Different rounds of insecticide nebulization were done in these three locations throughout the month. All these situations complicate our understanding of the normal course of the outbreak. Nonetheless, our BDSKY analysis evidenced this gap in sampling reducing the impact in our evolutionary/epidemiological inferences.

If we recall the SIRBD results (Figure 16) we can notice that the highest incidence of cases occurred around February and March and the figures of April represent mainly prevalence, indicating the inertia of the transmission wave. The posterior estimates of incidence obtained with Outbreaker (Figure 28) evidenced that the number of cases was very low around mid April, but also points out to a gap in the estimates around this time. If we actually follow the general tendency of the posterior counts we can infer that the number of cases were actually higher in this period. Under these assumptions the multimodal (twin-peak) profile would no reflect the true epidemics, being the deep centered around day 100, an artifact due to reduced sampling. Nevertheless, it is quite important to consider that after the application of insecticide nebulizations there was evidence of new transmission events in untreated areas and even at the treated areas, shown by green nodes in the transmission networks in Figure 23. Moreover, there was evidence for extended chains of transmission after the interference, indicated by the green sub-networks in Figure 21. Therefore a secondary outbreak of smaller proportions, (due possibly to control

measures and reduction of susceptibles), may have taken place in the wake of the main outbreak.

Figure 28- Posterior estimates of incidence using Outbreaker.

The graphic depicts the posterior incidence count values estimated during a MCMC run with outbreaker. The white line shows the observed tendency.

In summary, it is noteworthy to mention that is very difficult to control or eliminate *Ae. aegypti* mosquitoes because they have adaptations to the environment that make them highly resilient, or the ability to rapidly bounce back to initial population numbers after disturbances resulting from natural phenomena (e.g., droughts) or human interventions (e.g., control measures). Even when other factors may have contributed to the rapid reduction in the number of cases by the end of April, the decline followed the intense mosquito control actions in those transmission hubs that were identified in the current study as key for viral spread. The prompt response by the dengue surveillance team during outbreaks exemplifies how mapping infected patients and control measures in urban environments with these patients contribute to diminish dengue burden.

5 CONCLUSIONS

- (i) 525 samples were confirmed as being positive for dengue: 505 belonged to DENV-4 and the remainder to any other serotype. The isolation of the four serotypes evidences a rise in Brazilian urban hyperendemicity that constitutes a greater challenge for surveillance and control;
- (ii) All dengue samples from serotype 4 belonged to genotype 2A; the mean evolutionary rate estimated with BEAST approach was 7×10^{-6} (95% HPD interval 4.6×10^{-6} , 9.4×10^{-6}) substitutions per site per day.
- (iii) Two major DENV-4 lineages were identified; several imported cases were inferred with the Bayesian transmission tree approach. Altogether may be considered as an important warning for high levels of transmission in the region;
- (iv) The origin of the 2013 epidemic was inferred to started during the last two weeks of November 2012 (205 days back in the past 95%HPD interval 198, 212.5) when the municipality led the ranking of infestation of dengue in the State of São Paulo. Given so, surveillance systems need to prioritize early mosquito population control and monitor proactively the occurrence of initial low levels of transmission;
- (v) The neighborhoods Enseada and Pae Cará acted as major transmission foci. The prompt response by the dengue surveillance team, directing mosquito control to areas with higher number of NS1 + patient, appear to have had a considerable quenching effect on the outbreak. It exemplifies how mapping infected patients and strategically sound control measures directed toward their urban environments diminish dengue burden;
- (vi) The basic reproduction number for DENV-4 was around 1.7;

- (vii) We did not find any association between demographic or socioeconomic features and the incidence of dengue. Nonetheless, patients in a stage with high levels of NS1 and none immune response seems to be important for transmission.

REFERENCES*

- BARCELLOS, C.; PUSTAI, A. K.; WEBER, M. A.; BRITO, M. R. V. Identificação de locais com potencial de transmissão de dengue em Porto Alegre através de técnicas de geoprocessamento. **Rev. Soc. Bras. Med. Trop.**, v. 38, p. 246-250, 2005.
- BARRETO, F. R.; TEIXEIRA, M. G.; COSTA, M. D. C. N.; CARVALHO, M. S.; BARRETO, M. L. Spread pattern of the first dengue epidemic in the city of Salvador, Brazil. **BMC Public Health**, n. 8, p. 51, 2008.
- BASTOS, MDE. S.; FIGUEIREDO, R. M.; RAMASAWMY, R.; ITAPIREMA, E.; GIMAQUE J. B.; SANTOS, L. O.; FIGUEIREDO, L. T.; MOURÃO, M. P. Simultaneous circulation of all four dengue serotypes in Manaus, State of Amazonas, Brazil in 2011. **Rev. Soc. Bras. Med. Trop.**, v. 45, n. 3, p. 393-394, 2012.
- BECKETT, C. G.; KOSASIH, H.; FAISAL, I.; NURHAYATI; TAN, R.; WIDJAJA, S.; LISTIYANINGSIH, E.; MA'ROEF, C.; WURYADI, S.; BANGS, M. J.; SAMSI, T. K.; YUWONO D.; HAYES, C. G.; PORTER, K. R. Early detection of dengue infections using cluster sampling around index cases. **Am. J. Trop. Med. Hyg.**, v. 72, p. 777-782, 2005.
- BHATT, S.; GETHING, P. W.; BRADY, O. J.; MESSINA, J. P.; FARLOW, A. W.; MOYES, C. L.; DRAKE, J. M.; BROWNSTEIN, J. S.; HOEN, A. G.; SANKOH, O.; MYERS, M. F.; GEORGE, D. B.; JAENISCH, T.; WINT, G. R.; SIMMONS, C. P.; SCOTT, T. W.; FARRAR, J. J.; HAY, S. I. The global distribution and burden of dengue. **Nature**, v. 496, n. 7446, p. 504-507, 2013.
- BIELEJEC, F.; RAMBAUT, A.; SUCHARD, M. A.; LEMEY, P. SPREAD: Spatial Phylogenetic Reconstruction of Evolutionary Dynamics. **Bioinformatics**, v. 27, n. 20, p. 2910-2912, 2011.
- BOUCKAERT, R.; HELED, J.; KÜHNERT, D.; VAUGHAN, T.G.; WU, C-H.; XIE, D.; SUCHARD, M. A.; RAMBAUT, A.; DRUMMOND, A. J. BEAST2: A software platform for Bayesian evolutionary analysis. **PLoS Comput. Biol.**, v. 10, n. 4, p. e1003537, 2014.
- BRATHWAITE DICK, O.; SAN MARTÍN J. L.; MONTOYA, R. H.; DEL DIEGO, J.; ZAMBRANO, B.; DAYAN, G. H.; The history of dengue outbreaks in the Americas. **Am. J. Trop. Med. Hyg.**, v. 87, n. 4, p. 584-593, 2012.
- BURKE, D. S.; NISALAK, A.; JOHNSON D. E.; SCOTT, R. M. A prospective study of dengue infections in Bangkok. **Am. J. Trop. Med. Hyg.**, v. 38, p. 172-180, 1988
- CENTRO DE PESQUISAS METEOROLÓGICAS E CLIMÁTICAS APLICADAS À AGRICULTURA – CEPAGRI. Clima dos Municípios Paulistas. 2014. Available at: <<http://www.cpa.unicamp.br/outras-informacoes/clima-dos-municipios-paulistas.html>> accessed on: 05 Oct. 2014.

* De acordo com:

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: Informação e documentação: referências: elaboração. Rio de Janeiro, 2002.

CHAKRAVARTI, A.; ARORA, R.; LUXEMBURGER, C. Fifty years of dengue in India. **Trans. R. Soc. Trop. Med. Hyg.**, v. 106, p. 273–282, 2012.

CHAN, M.; JOHANSSON, M. A. The incubation periods of Dengue viruses. **PLoS One**, v. 7, n. 11, p. e50972, 2012.

CHEN, W. J.; CHEN, S. L.; CHIEN, L. J.; CHEN, C. C.; KING, C. C.; HARN, M. R.; HWANG, K. P.; FANG, J. H. Silent transmission of dengue virus in southern Taiwan. **Am. J. Trop. Med. Hyg.**, v. 55, p. 12-16, 1996.

CHEN, R.; VASILAKIS, N. Dengue--quo tu et quo vadis? **Viruses**, v. 3, n. 9, p. 1562-1608, 2011.

CHIARAVALLOTI NETO, F. C.; MORAES, M. S.; FERNANDES, M. A. Avaliação dos resultados de atividades de incentivo à participação da comunidade no controle da dengue em um bairro periférico do município de São José do Rio Preto, São Paulo, e da relação entre conhecimentos e práticas desta população. **Cad. Saude Publica**, v. 14, n. 2, p. 101-109, 1998.

CENTERS FOR DISEASE CONTROL AND PREVENTION. Dengue. 2014. Available at: <<http://www.cdc.gov/Dengue/epidemiology/index.html>>. Accessed in: 05 Oct. 2014.

COSTA, R. L.; VOLOCH, C. M.; SCHRAGO, C. G. Comparative evolutionary epidemiology of dengue virus serotypes. **Infect. Genet. Evol.**, v. 12, n. 2, p. 309-314, 2012.

COSTA, A. I. P.; NATAL, D. Distribuição espacial da dengue e determinantes socioeconômicos em localidade urbana no Sudeste do Brasil. **Rev. Saude Publica**, v. 32, p. 232-236, 1998.

CUONG, H. Q.; VU, N. T.; CAZELLES, B.; BONI, M. F.; THAI, K. T.; RABAA, M. A.; QUANG, L. C.; SIMMONS, C. P.; HUU, T. N.; ANDERS, K. L. Spatiotemporal dynamics of dengue epidemics, southern Vietnam. **Emerg. Infect. Dis.**, v. 19, n. 6, p. 945-953, 2013.

DARRIBA, D.; TABOADA, G. L.; DOALLO, R.; POSADA, D. jModelTest 2: more models, new heuristics and parallel computing. **Nat. Methods**, v. 9, n. 8, p. 772, 2012.

DRUMMOND, A. J.; HO, S. Y. W.; PHILLIPS, M. J.; RAMBAUT, A. Relaxed Phylogenetics and Dating with Confidence. **PLoS Biol.**, v. 4, n. e88, 2006.

EDGAR, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. **BMC Bioinformatics**, v. 5, p. 113, 2004.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Res.**, v. 32, n. 5, p. 1792-1797, 2004.

FAVIER, C.; DEGALLIER, N.; ROSA-FREITAS, M. G.; BOULANGER, J. P.; COSTA LIMA J. R.; LUITGARDS-MOURA, J. F.; MENKÈS, C. E.; MONDET, B.; OLIVEIRA, C.; WEIMANN, E. T.; TSOURIS, P. Early determination of the reproductive number for vector-borne diseases: the case of dengue in Brazil. **Trop. Med. Int. Health**, v. 11, n. 3, p. 332-340, 2006.

FOSTER, J. E.; BENNETT, S. N.; VAUGHAN, H.; VORNDAM, V.; MCMILLAN, W. O.; CARRINGTON, C. V. Molecular evolution and phylogeny of dengue type 4 virus in the Caribbean. **Virology**, v. 306, n. 1, p. 126-134, 2003.

FRAIHA, H. Reinfestação do Brasil pelo *Aedes aegypti*. Considerações sobre o risco de urbanização do vírus da febre amarela Silvestre na região reinfestada. **Rev. Inst. Med. Trop. São Paulo**, v. 10, n. 5, p. 289-294, 1968.

GEBHARD, L. G.; FILOMATORI, C. V.; GAMARNIK, A. V. Functional RNA elements in the dengue virus genome. **Viruses**, v. 3, n. 9, p. 1739-1756, 2011.

GIBBONS, R. V.; KALANAROOJ, S.; JARMAN, R. G.; NISALAK, A.; VAUGHN, D. W.; ENDY, T. P.; MAMMEN, M. P. JR.; SRIKIATKHACHORN, A. Analysis of repeat hospital admissions for dengue to estimate the frequency of third or fourth dengue infections resulting in admissions and dengue hemorrhagic fever, and serotype sequences. **Am. J. Trop. Med. Hyg.**, v. 77, n. 5, p. 910-913, 2007.

GUBLER, D. J. Dengue and dengue hemorrhagic fever. **Clin. Microbiol. Rev.**, v. 11, n. 3, p. 480-496, 1998.

GUBLER, D. J. Dengue/dengue haemorrhagic fever: history and current status. **Novartis Found Symp.**, v. 277, p. 3-22, 2006.

GUBLER, D. J. The economic burden of dengue. **Am. J. Trop. Med. Hyg.**, v. 86, n. 5, p. 743-774, 2012.

GUINDON, S.; GASCUEL, O. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. **Syst. Biol.**, v. 52, p. 696-704, 2003.

GUZMAN, M. G.; ALVAREZ, M.; HALSTEAD, S. B. Secondary infection as a risk factor for dengue hemorrhagic fever/dengue shock syndrome: an historical perspective and role of antibody-dependent enhancement of infection. **Arch. Virol.**, v. 158, n. 7, p. 1445-1459, 2013.

HALL, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symp. Ser.**, v. 41, p. 95-98, 1999.

HALSTEAD, S. B. Pathogenesis of dengue: challenges to molecular biology. **Science**, v. 239, p. 476-481, 1988.

HALSTEAD, S. B. Dengue virus-mosquito interactions. **Annu. Rev. Entomol.**, v. 53, p. 273-291, 2008.

HALSTEAD, S. B. Dengue. **Lancet**, v. 370, p. 1644-1652, 2007.

HOLMES, E. C. Patterns of intra- and interhost nonsynonymous variation reveal strong purifying selection in dengue virus. **J. Virol.**, v. 77, n. 20, p. 11296-11298, 2003.

HOLMES, E. C. The phylogeography of human viruses. **Mol. Ecol.**, v. 13, n. 4, p. 745-756, 2004.

INNIS, B. L. Dengue and dengue hemorrhagic fever. In: PORTERFIELD, J. S. **Exotic viral infections**. London: Chapman and Hall, 1995. p. 103–140.

JEEFOO, P.; TRIPATHI, N. K.; SOURIS, M. Spatio-temporal diffusion pattern and hotspot detection of dengue in Chachoengsao province, Thailand. **Int J Environ Res Public Health**, v. 8, n. 1, p. 51-74, 2011.

JOMBART, T.; CORI, A.; DIDELOT, X.; CAUCHEMEZ, S.; FRASER, C.; FERGUSON, N. Bayesian reconstruction of disease outbreaks by combining epidemiologic and genomic data. **PLoS Comput. Biol.**, v. 10, n. 1, p. e1003457, 2014.

KARABATSOS, N. **International catalogue of arboviruses**. San Antonio: American Society of Tropical Medicine and Hygiene, 1985.

KASSIM, F. M.; IZATI, M. N.; TGROGAYAH, T. A.; APANDI, Y. M.; SAAT, Z. Use of dengue NS1 antigen for early diagnosis of dengue virus infection. **Southeast Asian J. Trop. Med. Public Health**, v. 42, n. 3, p. 562-569, 2001.

KLUNGTHONG, C.; ZHANG, C.; MAMMEN, M. P. JR.; UBOL, S.; HOLMES, E. C. The molecular epidemiology of dengue virus serotype 4 in Bangkok, Thailand. **Virology**, v. 329, n. 1, p. 168-179, 2004.

KÜHNERT, D.; STADLER, T.; VAUGHAN, T. G.; DRUMMOND, A. J. Simultaneous reconstruction of evolutionary history and epidemiological dynamics from viral sequences with the birth-death SIR model. **J. R. Soc. Interface**, v. 11, n. 94, p. 20131106, 2014.

KUNO, G.; CHANG, G. J.; TSUCHIYA, K. R.; KARABATSOS, N.; CROPP, C. B. Phylogeny of the genus Flavivirus. **J. Virol.**, v. 72, n. 1, p. 73-83, 1998.

LANCIOTTI, R. S.; CALISHER, C. H.; GUBLER, D. J.; CHANG, G. J.; VORNDAM, A. V. Rapid detection and typing of dengue viruses from clinical samples by using reverse transcriptase-polymerase chain reaction. **J. Clin. Microbiol.**, v. 30, n. 3, p. 545-551, 1992.

LEMEY, P.; RAMBAUT, A.; DRUMMOND, A. J.; SUCHARD, M. A. Bayesian Phylogeography Finds Its Roots. **PLoS Comput. Biol.**, v. 5, p. e1000520, 2009.

LEMEY, P.; RAMBAUT, A.; WELCH, J. J.; SUCHARD, M. A. Phylogeography Takes a Relaxed Random Walk in Continuous Space and Time. **Mol. Biol. Evol.**, v. 27, p. 1877-1885, 2010.

LOBO, F. P.; MOTA, B. E.; PENA, S. D.; AZEVEDO, V.; MACEDO, A. M.; TAUCH, A.; MACHADO, C. R.; FRANCO, G. R. Virus-host coevolution: common patterns of nucleotide motif usage in Flaviviridae and their hosts. **PLoS One**, v. 4, n. 7, p. e6282, 2009.

MACHADO, J. P.; OLIVEIRA, R. M.; SOUZA-SANTOS, R. Spatial analysis of dengue occurrence and living conditions in Nova Iguaçu, Rio de Janeiro State, Brazil. **Cad. Saude Publica**, v. 25, n. 5, 2009.

MADDISON, W. P.; MADDISON, D. R. Mesquite: a modular system for evolutionary analysis. 2014. Available at: <<http://mesquiteproject.org> 2014>. Accessed in: 05 Oct. 2014.

MASSAD, E.; BURATTINI, M. N.; COUTINHO, F. A.; LOPEZ, L. F. Dengue and the risk of urban Yellow fever reintroduction in São Paulo State, Brazil. **Rev. Saude Publica**, v. 37, p. 477–484, 2003.

MEDRONHO, R. A. **Geoprocessamento e saúde: uma nova abordagem do espaço no processo saúde doença**. Rio de Janeiro: Editora Fiocruz, 1995.

MINISTÉRIO DA SAÚDE. **Reunião técnica do programa de combate à febre amarela e dengue**. Relatório final. Brasília: 1991.

MONDINI, A.; CHIARAVALLLOTI NETO, F. Variáveis socioeconômicas e a transmissão de dengue. **Rev. Saude Publica**, v. 41, p. 923-930, 2007.

MONDINI, A.; DE MORAES BRONZONI, R. V.; NUNES, S. H.; CHIARAVALLLOTI NETO, F.; MASSAD, E.; ALONSO, W. J.; LÁZZARO, E. S.; FERRAZ, A. A.; DE ANDRADE ZANOTTO, P. M.; NOGUEIRA, M. L. Spatio-temporal tracking and phylodynamics of an urban dengue 3 outbreak in São Paulo, Brazil. **PLOS Negl. Trop. Dis.**, v. 3, n. 5, p. e448, 2009.

NOGUEIRA, R. M.; MIAGOSTOVICH, M. P.; LAMPE, E.; SCHATZMAYR, H. G. Isolation of dengue virus type 2 in Rio de Janeiro. **Mem. Inst. Oswaldo Cruz**, v. 85, p. 253, 1990.

NOGUEIRA, R. M.; DE ARAÚJO, J. M.; SCHATZMAYR, H. G. Dengue viruses in Brazil, 1986-2006. **Rev. Panam. Salud Publica**, v. 22, n. 5, p. 358-363, 2007.

NORMILE, D. Surprising new dengue virus throws a spanner in disease control efforts. **Science**, v. 342, n. 6157, p. 415, 2013.

NUNES, M. R.; FARIA, N. R.; VASCONCELOS, H. B.; MEDEIROS, D. B.; SILVA DE LIMA, C. P.; CARVALHO, V. L.; PINTO DA SILVA, E. V.; CARDOSO, J. F.; SOUSA, E. C. JR; NUNES, K. N.; RODRIGUES, S. G.; ABECASIS, A. B.; SUCHARD, M. A.; LEMEY, P.; VASCONCELOS, P. F. Phylogeography of dengue virus serotype 4, Brazil, 2010-2011. **Emerg. Infect. Dis.** v. 18, n. 11, p. 1858-1864, 2012.

OSANAI, C. H. **A epidemia de dengue em Boa Vista, Território. Federal de Roraima, 1981-1982**. 1984. 127 f. Dissertação (Mestrado) - Escola Nacional de Saúde Pública, Rio de Janeiro, 1984.

Pan American Health Organization. 2001. Dengue in the Americas timeline. Available at: <http://www1.paho.org/English/AD/DPC/CD/dengue_finaltime.doc>. Accessed in: 05 Oct. 2014.

PAN AMERICAN HEALTH ORGANIZATION. Dengue 2014. Available at: <http://www.paho.org/hq/index.php?option=com_topics&view=article&id=1&Itemid=40734>. Accessed in: 05 Oct. 2014.

PIERSON, T. C. Program description of the laboratory of viral diseases from the National Institute of Allergy and Infectious Diseases. 2012. Available at: <http://www.niaid.nih.gov/labsandresources/labs/aboutlabs/lvd/viralpathogenesissection/Pages/default.aspx>. Accessed in: 05 Oct. 2014.

PINHO, S. T.; FERREIRA, C. P.; ESTEVA, L.; BARRETO, F. R.; MORATO, E.; SILVA, V. C.; TEIXEIRA, M. G. Modelling the dynamics of dengue real epidemics. **Philos. Trans. A Math. Phys. Eng. Sci.**, v. 368, n. 1933, p. 5679-5693, 2010.

PONTES, R. J. S.; FABRO, A. L. D.; ROCHA, G. M.; SANTIAGO, R. C.; FIGUEIREDO, L. T. M.; SILVA, A. A. M. C.; GAROTTI, V. D. O.; PINTYA, J. M. P. Epidemia de dengue em Ribeirão Preto, SP, Brasil: nota prévia. **Rev. Saude Publica**, v. 25, p. 315-317, 1991.

PONTES, R. J. S. **Estudo da epidemia de dengue no município de Ribeirão Preto- SP, 1990-1991**. 1992. Tese (Doutorado) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, São Paulo, 1992.

PONTES, R. J. S.; RUFFINO-NETTO, A. Dengue in urban locality of Southeastern, Brazil: epidemiological aspects. **Rev. Saude Publica**, v. 28, n. 3, p. 218-227, 1994.

POON, A. F. Y.; FROST, S. D. W.; POND, S. L. K. Detecting selection in DNA sequences with Datamonkey. In: POSADA, D. **Bioinformatics for DNA sequence analysis**. New Jersey: Humana Press, 2009.

RAGHWANI, J.; RAMBAUT, A.; HOLMES, E. C.; HANG, V. T.; HIEN, T. T.; FARRAR, J.; WILLS, B.; LENNON, N. J.; BIRREN, B. W.; HENN, M. R.; SIMMONS, C. P. Endemic dengue associated with the co-circulation of multiple viral lineages and localized density-dependent transmission. **PLoS Pathog.** v. 7, n. 6, p. e1002064, 2011.

RASMUSSEN, D. A.; BONI, M. F.; KOELLE, K. Reconciling phylodynamics with epidemiology: the case of dengue virus in southern Vietnam. **Mol. Biol. Evol.**, v. 31, n. 2, p. 258-271, 2014.

RICO-HESSE, R. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. **Virology**, v. 174, n. 2, p. 479-493, 1990.

RIGAU-PÉREZ, J. G. The early use of break-bone fever (Quebranta huesos, 1771) and dengue (1801) in Spanish. **Am. J. Trop. Med. Hyg.**, v. 59, n. 2, p. 272-274, 1998.

ROCCO, I. M.; FERREIRA, I. V.; KATZ, G.; SOUZA, L. T. M.; KIMURA-GUSHIKEN, E. K.; MENDES, K. H. C.; BASSI, M. G.; DEL GUERCIO, V. M. F.; TENGAN, C. H.; GALIMBERTTI, M. Z.; KAVAKAMA, B. B. Ocorrência de Dengue no estado de São Paulo, Brasil, de 1986 a 1996. **Rev. Inst. Adolfo Lutz**, v. 57, n. 1, p. 7-12, 1998.

RODRIGUEZ-BARRAQUER, I.; CORDEIRO, M. T.; BRAGA, C.; SOUZA, W. V.; MARQUES, E. T.; CUMMINGS, D. A. From Re-Emergence to Hyperendemicity: The Natural History of the Dengue Epidemic in Brazil. **PLOS Negl. Trop. Dis.**, v. 5, p. e935, 2011.

RUSH, B. An accounts of the Bilious Remitting Fever. **Am. J. Med.**, reprinted from Medical Inquiries and Observations, v. 11, 1809.

SABIN, A. B. Research on dengue during World War II. **Am. J. Trop. Med. Hyg.**, v. 1, n. 1, p. 30-50, 1952.

SCHATZMAYR, H. G.; NOGUEIRA, R. M.; TRAVASSOS DA ROSA, A. P. An outbreak of dengue virus at Rio de Janeiro. **Mem. Inst. Oswaldo Cruz**, v. 81, p. 245-246, 1986.

SCHREIBER, M. J.; HOLMES, E. C.; ONG, S. H.; SOH, H. S.; LIU, W.; TANNER, L.; AW, P. P.; TAN, H. C.; NG, L. C.; LEO, Y. S.; LOW, J. G.; ONG, A.; OOI, E. E.; VASUDEVAN, S. G.; HIBBERD, M. L. Genomic epidemiology of a dengue virus epidemic in urban Singapore. **J. Virol.**, v. 83, n. 9, p. 4163-4173, 2009.

STADLER, T.; KÜHNERT, D.; BONHOEFFER, S.; DRUMMOND, A. J. Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). **Proc. Natl. Acad. Sci. U.S.A.**, v. 110, n. 1, p. 228-233, 2013.

SECRETARIA DE ESTADO DA SAÚDE DE SÃO PAULO. Superintendência de Controle de Endemias. **Plano emergencial de controle do Aedes aegypti para o verão de 1989**. São Paulo, 1989.

SECRETARIA DE ESTADO DA SAÚDE DE SÃO PAULO. Superintendência de Controle de Endemias. **Programas de controle de vetores**. São Paulo, 1991a.

SECRETARIA DE ESTADO DA SAÚDE DE SÃO PAULO. Superintendência de Controle de Endemias. **Proposta de descentralização do controle de endemias**. São Paulo, 1991b.

SECRETARIA DE ESTADO DA SAÚDE DE SÃO PAULO. Superintendência de Controle de Endemias. **Vigilância epidemiológica do dengue no Estado de São Paulo**, 1991c.

SECRETARIA DE ESTADO DA SAÚDE DE SÃO PAULO. Superintendência de Controle de Endemias. **Plano emergencial para o controle dos vetores do dengue e da febre amarela no verão de 1991-1992**. São Paulo, 1991d.

SHEPARD, D. S.; COUDEVILLE, L.; HALASA, Y. A.; ZAMBRANO, B.; DAYAN, G. H. Economic impact of dengue illness in the Americas. **Am. J. Trop. Med. Hyg.**, v. 84, n. 2, p. 200-207, 2011.

SUAYA, J. A.; SHEPARD, D. S.; SIQUEIRA, J. B.; MARTELLI, C. T.; LUM, L. C.; TAN, L. H.; KONGSIN, S.; JIAMTON, S.; GARRIDO, F.; MONTROYA, R.; ARMIEN, B.; HUY, R.; CASTILLO, L.; CARAM, M.; SAH, B. K.; SUGHAYYAR, R.; TYO, K. R.; HALSTEAD, S. B. Cost of dengue cases in

eight countries in the Americas and Asia: a prospective study. **Am. J. Trop. Med. Hyg.**, v. 80, n. 5, p. 846-855, 2009.

SUZUKI, Y.; GOJOBORI, T. A method for detecting positive selection at single amino acid sites. **Mol. Biol. Evol.**, v. 16, p. 1315–1328, 1999.

TEIXEIRA, M. G.; COSTA, M. C. N.; COELHO, G.; BARRETO, L. M. Recent shifts in age pattern of Dengue Hemorrhagic fever, Brazil. **Emerging Infect. Dis.**, v. 14, p. 1663, 2008.

TEIXEIRA, M. G.; SIQUEIRA, J. B. JR; FERREIRA, G. L.; BRICKS, L.; JOINT, G. Epidemiological trends of dengue disease in Brazil (2000-2010): a systematic literature search and analysis. **PLOS Negl. Trop. Dis.**, v. 7, n. 12, p. e2520, 2013.

TEIXEIRA, MDA. G.; BARRETO, M. L.; COSTA, MDA. C.; FERREIRA, L. D.; VASCONCELOS, P. F.; CAIRNCROSS, S. **Trop. Med. Int. Health.**, v. 7, n. 9, p. 757-762, 2002.

TEIXEIRA, T. R.; CRUZ, O. G. Spatial modeling of dengue and socio-environmental indicators in the city of Rio de Janeiro, Brazil. **Cad Saude Publica.**, v. 27, n. 3, p. 591-602, 2011.

TWIDDY, S. S.; HOLMES, E. C.; RAMBAUT, A. Inferring the rate and time-scale of dengue virus evolution. **Mol. Biol. Evol.**, v. 20, n. 1, p. 122-129, 2003.

VASCONCELOS, P. F.; LIMA, J. W.; RAPOSO, M. L.; RODRIGUES, S. G.; ROSA, J. F. S. T.; AMORIM, S. M. C.; ROSA E. S. T.; MOURA, C. M. P.; FONSECA, N.; ROSA, A. P. A. T. Inquérito sorológico na Ilha de São Luiz durante uma epidemia de dengue no Maranhão. **Rev. Soc. Bras. Med. Trop.**, v. 32, p. 171-179, 1999.

VASILAKIS, N.; CARDOSA, J.; HANLEY, K. A.; HOLMES, E. C.; WEAVER, S. C. Fever from the forest: prospects for the continued emergence of sylvatic dengue virus and its impact on public health. **Nat. Rev. Microbiol.**, v. 9, n. 7, p. 532-541, 2011.

VAZQUEZ-PROKOPEC, G. M.; KITRON, U.; MONTGOMERY, B.; HORNE, P.; RITCHIE, S. A. Quantifying the spatial dimension of dengue virus epidemic spread within a tropical urban environment. **PLOS Negl. Trop. Dis.** v. 4, n. 12, p. e920, 2010.

VILLABONA-ARENAS, C. J.; ZANOTTO, P. M. Evolutionary history of Dengue virus type 4: insights into genotype phylodynamics. **Infect. Genet. Evol.**, v. 11, n. 5, p. 878-875, 2011.

VILLABONA-ARENAS, C. J.; Botelho, A. V.; Botelho, A. C.; Passos, S. D.; Zanotto, P. M. The burden of dengue: Jundiaí, Brazil - January 2010. **Rev. Assoc. Med. Bras.**, v. 58, n. 4, 2012.

WANG, E.; NI, H.; XU, R.; BARRETT, A. D.; WATOWICH, S. J.; GUBLER, D. J.; WEAVER, S. C. Evolutionary relationships of endemic/epidemic and sylvatic dengue viruses. **J. Virol.**, v. 74, n. 7, p. 3227-3234, 2000.

WETTSTEIN, Z. S.; FLEMING, M.; CHANG, A. Y.; COPENHAVER, D. J.; WATESKA, A. R.; BARTSCH, S. M.; LEE, B. Y.; KULKARNI, R. P. Total economic cost and burden of dengue in Nicaragua: 1996-2010. **Am. J. Trop. Med. Hyg.**, v. 87, n. 4, p. 616-622, 2012.

WHITEHEAD, S. S.; BLANEY, J. E.; DURBIN, A. P.; MURPHY, B. R. Prospects for a dengue virus vaccine. **Nat. Rev. Microbiol.**, v. 5, n. 7, p. 518-528, 2007.

WORLD HEALTH ORGANIZATION. Dengue 2014. Available at: <<http://www.who.int/topics/dengue/en/>>. Accessed in: 05 Oct. 2014

AKIFUMI, Y.; TADAHIRO, S.; TAKESHI, K.; TERUO, Y.; KAZUYOSHI, I. Origin and distribution of divergent dengue virus: novel database construction and phylogenetic analyses. **Future Virol.**, v. 8, n. 11, p. 1061-1083, 2013.

YAP, G.; LI, C.; MUTALIB, A.; LAI, Y. L.; NG, L. C. High rates of inapparent dengue in older adults in Singapore. **Am. J. Trop. Med. Hyg.**, v. 88, n. 6, p. 1065-1069, 2013.

YEW, Y. W.; YE, T.; ANG, L. W.; NG, L. C.; YAP, G.; JAMES, L.; CHEW, S. K.; GOH, K. T. Seroepidemiology of dengue virus infection among adults in Singapore. **Ann. Acad. Med. Singap.**, v. 38, n. 8, p. 667-675, 2009.

ZHANG, C.; MAMMEN, M. P. JR; CHINNAWIROTPISAN, P.; KLUNGTHONG, C.; RODPRADIT, P.; NISALAK, A.; VAUGHN, D. W.; NIMMANNITYA, S.; KALAYANAROOJ, S.; HOLMES, E. C. Structure and age of genetic diversity of dengue virus type 2 in Thailand. **J. Gen. Virol.**, v. 87, n. 4, p. 873-883, 2006.

ZWICKL, D. J. **Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion**. 2006. [Ph.D. thesis] – The University of Texas at Austin, Austin, 2006

APPENDIX A – Manuscripts

Evolutionary history of Dengue virus type 4: Insights into genotype phylodynamics

Christian Julián Villabona-Arenas, Paolo Marinho de Andrade Zanotto *

Laboratory of Molecular Evolution and Bioinformatics, Department of Microbiology, Biomedical Sciences Institute, University of São Paulo, CEP: 05508-000, Av. Prof. Lineu Prestes, 1734 São Paulo, Brazil

ARTICLE INFO

Article history:

Received 4 November 2010
Received in revised form 2 February 2011
Accepted 9 February 2011
Available online 16 February 2011

Keywords:

Dengue virus type 4
Evolution
Phylogeography
Phylodynamics

ABSTRACT

Dengue virus type 4 (DENV-4) circulates in tropical and subtropical countries from Asia and the Americas. Despite the importance of dengue virus distribution, little is known about the worldwide viral spread. Following a Bayesian phylogenetic approach we inferred the evolutionary history of 310 isolates sampled from 37 countries during the time period 1956–2008 and the spreading dynamics for genotypes I and II. The region (tropical rainforest biome) comprised by Malaysia–Thailand was the most likely ancestral area from which the serotype has originated and spread. Interestingly, cross-correlation analysis on demographic time series with the Asian sequences showed a statistically significant negative correlation that could be suggestive of competition among genotypes within the same serotype.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Four immunologically related, single positive-stranded RNA viruses known as dengue serotypes (DENV-1 through DENV-4) are responsible for causing dengue fever (DF) and dengue haemorrhagic fever (DHF) (Gubler, 1998). Global incidence of dengue has grown dramatically in recent years and up to epidemiological week number 30 of 2010, 1,082,402 cases have been reported in the region of the Americas (Pan American Health Organization, 2010).

Dengue viruses are geographically widespread in tropical and sub-tropical climates, mostly in urban and semi-urban areas, and largely determined by the range and dispersal of their main vectors *Aedes aegypti* and *A. albopictus*. Besides, modern transportation provides an efficient mechanism to quickly move dengue viruses and mosquitoes around the globe (Wilder-Smith and Gubler, 2008). Consequently, one primary way to prevent dengue virus transmission is to combat these disease-carrying mosquitoes. A significant cause of dengue emergence in Latin America involved the reintroduction of *A. aegypti* following cessation of mosquito control programs, which had been in place until the early 1970s (Gubler, 1998; San Martín et al., 2010). Because antibody-dependent enhancement has been described for dengue viruses (Halstead et al., 1970; Halstead and O'Rourke, 1977; Halstead, 2009; Dejnirattisai et al., 2010), DHF cases occur regularly in places where two or more different serotypes are sequentially circulating.

Epidemiological studies have shown an increased risk of developing DHF after a secondary dengue infection (Guzman et al., 1990; Kliks et al., 1988; Sangkawibha et al., 1984; Thomas et al., 2008) highlighting the urgent need for a greater understanding of virus spreading to reverse the epidemiological trend.

With the climate change and the urban environment (e.g. increases in the size and population density on cities, irregular water supply and management and poor health infrastructures services) influencing the geographical distribution of *Aedes* sp. (Jansen and Beebe, 2010), the unabated increase of dengue, has called for research focused on the introduction routes of each serotype (Rico-Hesse, 1990; Lanciotti et al., 1997; Messer et al., 2003). Nevertheless, despite the importance of dengue virus distribution, little is known about the rates and patterns of global viral spread. Furthermore, within each serotype there are clusters of lineages (i.e., genotypes), which often have spatial and temporal disjunct distributions (Holmes, 2004). Recently, Ty Hang et al. (2010) found significant differences in viraemia levels among two different DENV-2 genotypes, which might explain an episode of genotype replacement in Vietnam, and emphasized that an understanding of DENV evolutionary dynamics can inform the development of vaccines and anti-viral drugs. In addition, Lourenço and Recker (2010) investigated some of the epidemiological factors that might determine the success and invasion dynamics of an advantageous dengue genotype.

Araújo et al. (2008) looked upon the global phylogeography and evolutionary history of the main DENV-3 genotypes using a parsimony approach and revealed geographical subdivision of populations in several country-specific clades. Here, we extend the

* Corresponding author. Tel.: +55 11 3091 8453; fax: +55 11 3091 7354.
E-mail address: pzanotto@usp.br (Zanotto).

study to DENV-4 and reconstruct the evolutionary relationships and the spatial dispersion patterns of worldwide representative DENV-4 strains using a Bayesian framework for inference of phylogeographic history.

2. Materials and methods

2.1. DENV-4 molecular data

Complete envelope gene sequences with known date and location of sampling which represent the available genetic diversity for DENV-4, were collected from GenBank. This resulted in a final data set of 310 sequences from 37 countries isolated during 1956–2008. [Supplementary Table 1](#) shows the number, date of isolation, and origin of the included isolates. Nucleotide sequences were aligned using Muscle 3.8.31 program ([Edgar, 2004](#)), and later manually edited with the Se-Al v2.0a11 program.

2.2. Phylodynamics of DENV-4

To reconstruct the evolutionary history for the complete data set of DENV-4 envelope sequences, the obtained alignment was used in a maximum likelihood (ML) optimization using Garli build 0.951 ([Zwickl, 2006](#)) that uses the General Time Reversible (GTR) ([Rodríguez et al., 1990](#)) model of nucleotide substitution and simultaneously optimizes topology, branch lengths, rate heterogeneity and model parameters. Because we had dates of sampling for each sequence, a maximum clade credibility (MCC) tree with dated tips and internal nodes was inferred using a Markov Chain Monte Carlo (MCMC) Bayesian approach under GTR model with Gamma-distributed rate variation (γ) and a proportion of invariable sites (I) using a relaxed (uncorrelated lognormal) molecular clock ([Drummond et al., 2006](#)) in the program BEAST v1.5.4 ([Drummond and Rambaut, 2007](#)). The GTR + γ + I substitution model was used since it was the best model obtained in jModelTest ([Posada, 2008](#)). We used a relaxed molecular clock because previous Envelope gene sequences analyses have suggest that dengue evolution generally approximates a molecular clock, but some minor rate differences occur ([Twiddy et al., 2003](#)). Four independent MCMC runs of four chains each were run for 50 million generations. Moreover, we estimated the evolutionary rate, the time of the most recent common ancestor (TMRCA) and the compounded parameter (*Neg*) assumed as the effective number of infections in time that correlates with viral population sizes under a Bayesian skyline coalescent tree prior ([Drummond et al., 2005](#)). Convergence of parameters during MCMC runs were assessed by their Effective Sample Size (ESS) reaching values above 200, as calculated with the TRACER v1.5 program ([Rambaut and Drummond, 2007](#)).

2.3. Phylogeography of DENV-4

To investigate the spread pattern of DENV-4 the country of origin of each sequence was used as a set of terminal unordered character states, represented by two capital letters using the ISO 3166-1 alpha-2 code (http://www.iso.org/iso/english_country_names_and_code_elements). Most reconstruction methods of the history of pathogens use parsimony approaches that reconstruct a minimum set of migration events consistent with the observed phylogeny ([Pybus and Rambaut, 2009](#); [Parker et al., 2008](#)). We used a time-scaled phylogeny to reconstruct the DENV-4 spread patterns using a standard continuous-time Markov chain (CTMC) model with the Bayesian stochastic search variable selection (BSSVS) procedure in BEAST ([Lemey et al., 2009](#)). The description of geographic spread was obtained by MCMC sampling from the plausible set of trees using the procedures outlined above and

taking the inferred internal states as surrogates for migration events with a Bayes Factors above 25 that we considered as credible support.

2.4. Cross-correlation among DENV-4 genotypes

To estimate the degree to which genotypes I and II are correlated in abundance over time, we did cross-correlation analysis on *Neg* time series with the Asian sequences only. The use of the *Neg* time series obtained from the BSL is justified because the coalescent model recovers past cladogenetic events on viral genealogies that inform on changes in population size. It is important to notice that this demographic information is not available from data on counts of samples through time. Data from the Bayesian skyline on *Neg* over time generated with BEAST were put on a common time scale using the polynomial interpolation method implemented in Origin 8 (OriginLab, Northampton, MA). To establish if a relationship exists between both genotypes, we computed cross-correlation coefficients between the two synchronous series over a range of time lags from one up to ten years, corresponding to time spans between seasonally spaced cycles to minimize possible autocorrelation noise from data interpolation ([Chatfield, 1975](#)). Cross-correlations were done using SPSS v. 11.0.4 (SPSS for Mac Rel. 11.0.4.2005, SPSS Inc., Chicago).

3. Results and discussion

3.1. Estimating a rate of change

Previous analyses revealed that DENV-4 strains exhibited greater nucleotide sequence conservation than the other DENV serotypes (92%) and 96–100% conservation in envelope protein amino acids as reviewed by [Weaver and Vasilakis \(2009\)](#). We found a conservation of 85% in nucleotide sequences (89% when excluding sylvatic strains) and 95% in amino acids sequences of the envelope gene region. The substitution rate estimated for our complete DENV-4 data using a relaxed molecular clock model was 9.07×10^{-4} subs/site/year (HPD 7.57×10^{-4} to 1.06×10^{-3}). This estimated rate HPD includes that previously obtained by [Twiddy et al. \(2003\)](#) (6.02×10^{-4} subs/site/year; HPD 4.29×10^{-4} to 8.07×10^{-4} subs/site/year) using only 23 DENV-4 sequences.

3.2. A chronology of the global spread of DENV-4

Our analyses recognized two DENV-4 genotypes separating (I) Filipinos, Thai and Sri Lankan viruses from (II) Indonesian, French Polynesian, and Latin American ones ([Lanciotti et al., 1997](#)). Current DENV-4 phylogenies delineated two other minor genotypes: (III) the 1990s Thai strains that are distinct from other Thai isolates and (IV) the sylvatic isolates ([Klungthong et al., 2004](#)). Our phylogenetic analysis of worldwide 310 DENV-4 envelope gene sequences identified the previously described four genotypes, each with a posterior probability (BPP) of 1 ([Fig. 1](#)). We also estimated rates and the time of the most recent common ancestor (TMRCA) for each genotype. Our results confirmed that three DENV-4 human genotypes appear to have diverged from the currently available sylvatic strains around 157 years before present (YBP) (see the attached uncertainty for the estimates in [Table 1](#)) as previously suggested by [Twiddy et al. \(2003\)](#) (195.32 years; HPD 152.25–264.80). Subsequently, genotype III diverged in Asia (possibly Thailand) around 102 YBP from the ancestral of genotypes I and II, which split 84 YBP.

Genotype II has been circulating in the Americas since its introduction into the Caribbean in the 1980s from Southeast Asia ([Foster et al., 2003](#)). Interestingly, our MCC tree dates the event to 1978, in almost perfect agreement with the known history of this

Fig. 1. Maximum clade credibility (MCC) tree showing the phylogenetic relationships among DENV-4 genotypes, show as collapsed clades for clarity. The posterior probabilities and the time of the most recent common ancestor (TMRCA) are shown for main nodes. Light grey text represents available sequences by country (following the ISO 3166-1 alpha-2 code) included in each genotype. Clades size proportional to the number of sequences included. A full, detailed tree file is available as [Supplementary Data](#).

serotype. Furthermore, genotype II could be divided into two subclusters, each with a BPP of 1 and a between group mean distance of 6% nucleotide divergence. Noticeably, there was a geographical and temporal subdivision of strains, since subcluster I contained all the American isolates and most of the Asian isolates before year 2000 and subcluster II contained all Asian isolates after year 2000 and three older ones (both Indonesia 1973 and China 1978).

The ML tree recovered with Garli (data not shown) and the dated MCC tree obtained with Beast were congruent ([Full detailed trees available as Supplementary Data](#)). Therefore, we argue that the time stamping of samples that allows BEAST to adjust branch lengths according to dates of sampling did not appear to impact the topology, suggesting that our main conclusions were quite robust to method of inference. Both topologies were characterized by a tree in which the oldest sampled viruses clustered basally to those sampled more recently resulting in a pattern of evolution better defined by time period than by country-specific clades. A similar finding was reported for DENV-4 in the Caribbean by [Foster et al. \(2003\)](#), showing marked evolutionary shifts. In the Caribbean, the

most common genotype at any given time and place seems to arise from the same previously rare variant ([Bennett et al., 2003](#)).

3.3. Migration patterns of DENV-4

Our reconstruction of discrete ancestral states ([Fig. 1](#)) had the higher BPP values for both Malaysia and Thailand (data not shown). This agrees with the notion that a shared tropical rainforest biome region comprising these neighbouring countries could be the most probable place of origin for the spread of the serotype. Furthermore, it has been postulated that the greater diversity of sylvatic DENV in Malaysia suggests that an ancestor of all DENV arose in that area and diverged in the relatively distant past into the four serotypes recognized today ([Wang et al., 2000; Weaver and Vasilakis, 2009](#)). We would like to highlight here that all conclusions drawn on migration patterns are a function of the available sample of sequences. For the basal human node a BPP value of 79 was obtained likely meaning that the place of origin of the human DENV-4 samples analyzed cannot be resolved.

Table 1

Estimated DENV-4 substitutions rates and the time of the most recent common ancestor (TMRCA) estimated under a Bayesian skyline coalescent tree prior.

Set	Number of sequences	Rate (subs/site/year)	Rate 95% HPD (subs/site/year)	Node age	95% HPD (node age)	TMRCA (year)
DENV-4	310	9.07×10^{-4}	7.57×10^{-4} , 1.06×10^{-3}	157.27	82.77–240.89	1851
Genotype I	73	9.03×10^{-4}	1.80×10^{-4} , 1.95×10^{-3}	77.60	64.70–91.05	1930
Genotype II	228	6.17×10^{-4}	1.95×10^{-4} , 1.14×10^{-3}	59.09	47.93–72.10	1949
Genotype III	5	8.01×10^{-4}	3.81×10^{-4} , 1.25×10^{-3}	13.20	11.60–15.09	1995
Sylvatic	4	1.62×10^{-3}	2.54×10^{-4} , 3.70×10^{-3}	36.53	35.24–38.23	1971

Fig. 2. Plausible places of origin as inferred by ancestral state reconstructions and migratory routes of (A) DENV-4 genotypes I, (B) genotype II subclade II and (C) genotype II subclade I.

Nevertheless, in most instances Thailand is the most frequent assignment for the ancestral character state reconstructed. For genotype I (Fig. 2A), the most probable ancestral location was Thailand. The data suggested spreading from this country to the Philippines and traffic afterwards among the neighbouring countries and movement to Sri Lanka and India. For both subclusters of genotype II, the most probable ancestral location was Indonesia. In the Asian subclusters, there was circulation among regions in the Southeast reaching Southwest Pacific (Fig. 2B). Southeast Asia as an ancestral area would explain several DENV-4 outbreaks during 2002 in north Queensland, Australia, that have been traced to Thailand or Indonesia as their likely origin (Hanna et al., 2003). Moreover, in the case of subcluster I of genotype II there are an initial spread into French Polynesia, followed by spread to the Lesser Antilles (Fig. 2C). This finding corroborates the original suggestion, using fingerprinting data, that Caribbean DENV-4 viruses were related to viruses isolated from the Pacific islands one year before the onset of Caribbean epidemics (Henchal et al., 1986). Starting from Dominica and the French islands of Saint Barthelemy and Saint Martin, this genotype has rapidly spread all over other Caribbean localities and several other American countries (Centers for Disease Control and Prevention, 1981). Interestingly, our inference of phylogeographic history suggested the spread of DENV-4 from the American mainland back to the Southwest Pacific. However, we cannot accurately pinpoint its geographical origin due to the low BPP value obtained (0.16 grouping with El Salvador). The pattern of spread through time is shown in an animation (Supp.DENV4.kml) available as supplementary data to be visualized in Google Earth.

Carrington et al. (2005) have proposed that DENV-4 may have first entered South America, subsequently moving into the Caribbean. Here, we used a Bayesian stochastic search variable

selection (BSSVS) model, which enables us to reconstruct and test alternatives for the most parsimonious description of the phylogeographic diffusion process. Our results support the hypothesis that the initial movement tended to be from the Antilles followed by a bidirectional North-South spread into different regions of the American mainland. To further confirm this trend, we also used the *BayesMultistate* program (Pagel et al., 2004) from the Bayestraits software package (available at <http://www.evolution.rdg.ac.uk/>) to reconstruct under maximum likelihood how the major geographical areas (coded as discrete multistate traits: *Asia mainland*, *Southeast Asian islands/Southwest Pacific*, *Caribbean*, *Central-America*, and *South-America*) changed along a representative sample of trees recovered during the stationary phase of the MCMC search. The inferred rates of change among character states were taken as surrogates for migration rates among areas. Highest values (see Supplementary Table 2) were found for *Central-America* to *South-America* and the other way around, reflecting high viral traffic between both regions, followed by *Caribbean* to *South-America* and *Southeast Asian islands/Southwest Pacific* to *Asia mainland*, demonstrating both, the rapid spread of genotype II after arriving to the Antilles and the continued circulation of virus in Asia, respectively.

Countries such as Colombia, Ecuador, French Guiana, Peru, Suriname and Venezuela have regular DENV-4 outbreaks assumed to involve genotype II, since no other genotype has been detected in the Americas, according to Forshey et al. (2009) and reports of dengue cases from Pan American Health Organization (2010). However, the introduction of genotype I has been documented in the Brazilian Amazon (Figueiredo et al., 2008) possibly coming from an Asian country (Melo et al., 2009). Other than this latter report and a brief incursion of genotype II in Rondonia in 1982 (Recorded in GenBank), there is no further record of serotype 4

Fig. 3. Changes of the effective population size ($N_e.g$) estimated over time for all sequences from genotype I and genotype II as shown Bayesian skyline (BSL) plots in (A). Plotted below in (B) are the numbers of all sequences from which the BSL were inferred.

circulation in Brazil. Using the available sequences for Latin America, we did not find any evidence of the introduction of any other Asian DENV-4 genotype into the Americas, other than genotype II in 1978. Likewise, recent studies with DENV-3 have shown different genotypes co-circulating during outbreaks in Brazil (Aquino et al., 2009; Barcelos Figueiredo et al., 2008). Interestingly, DENV-3 genotype V had low prevalence for years (Aquino et al., 2009), and DENV-3 genotype I was associated with DHF cases in recent outbreaks (Barcelos Figueiredo et al., 2008).

3.4. The phylodynamics of DENV-4

Using the Bayesian skyline coalescent tree prior, we inferred the epidemic dynamics of DENV-4 through time. Fig. 3A shows a Bayesian skyline (BSL) plot on top of a histogram showing the number of sequences available every year from 1968 to 2008 for genotypes I and II (Fig. 3B). The demographic pattern for genotype I (black line) shows fluctuating dynamics characterized by two major peaks around 1981 and 2004 until 2008. On the other hand,

Fig. 4. Effective population size estimates (N_{eg}) estimated over time for Asian genotype I and genotype II (A), and the number of available sequences (C) are shown along with their respective correlograms (B) and (D) showing negative correlation coefficient values as a function of the lag size. The negative correlation pattern observed in the BSL (B) is not explained by changes of sampling density over time shown in (D), but by the reconstructed cladogenetic events in time (i.e., phylodynamics).

genotype II dynamics (light grey line) has been thoroughly characterized by Carrington et al. (2005) and Bennett et al. (2010). Briefly, there is a maintenance phase where N_{eg} values remain approximately constant followed by a growth phase after 1981, reflecting the introduction of DENV-4 into the Americas. Subsequently, an increased growth phase started around 1990 with minor fluctuations that appear to associate with the peaks in the number of samples by country. We argue that the increase in available samples does not explain the differences in growth trends observed between the BSLs. Because the BSLs in Fig. 3A depict the global growth of DENV-4, we sought to investigate the relationship among genotypes I and II in Asia, where both co-circulate. Interestingly, even after taking out the American sequences, the remaining Asian genotype I and genotype II sequences had BSLs (Fig. 4A) that were negatively correlated (significantly up to ten years apart, with a correlation coefficient of -0.731 and p -value <0.05) from 1976 to 2001 and became positively correlated

since 2001 (Fig. 4B). On the other hand, the same time series analysis done on samples over time (Fig. 4C) did not show any significant correlation between the genotypes that could explain the observed negative correlation between the BSLs (Fig. 4D). Because the Asian genotype I and genotype II lineages we sampled from were possibly directly subjected to cross-immunity while infecting human populations (either simultaneously or in succession), we argue that the significant negative correlation signal on the BSLs could reveal competition between the genotypes. Crucially, viral traits such as virulence and vector competence may be selected for in both humans and mosquitoes, causing potential differences in growth potential among lineages (Cologna et al., 2005). Moreover intra-serotypic competition could also help explain why at any given region, increase in a given serotype prevalence has been associated with an increase in its lineage diversity and a concomitant decrease in other serotype prevalence and diversity (i.e., clade replacement) (Zhang et al., 2005; Adams et al., 2006).

This study represents the largest phylogeographic analysis reported for DENV-4 to date. Our analysis evidenced the spatial dispersion patterns of the two main DENV-4 genotypes, and supports the notion that Thailand, Malaysia and Indonesia were sources of dengue lineages that have subsequently spread to other countries and continents. Moreover, while genotype I circulates in the Southeast Asia, viral strains of genotype II have evolved independently in the Americas and Southeast Asia for almost 30 years, with limited genetic exchange between the two land masses. This isolation may explain the diversification of viral lineages reflected by the accumulation of geographically structured polymorphisms (*i.e.*, regional synapomorphies). We also present some evidence suggestive of competition among genotypes within the same serotype. This relevant finding brings into question possible intra-serotype interactions in the ADE context. Further studies on the phylogeographic structure of dengue viruses are critical given that clinical severity has been shown to depend on the previous exposure to genetically distinct strains (Dejnirattisai *et al.*, 2010). Additional time-series analysis, with more data could help clarify if there are differences in inter-serotype cross-immunity in epidemic settings where different genotypes of distinct serotypes are present.

Acknowledgments

This work was supported by a FAPESP grant (2008/10713-6). CJVA is supported by a CAPES scholarship. PMAZ holds a CNPq – PQ scholarship.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.meegid.2011.02.007.

References

- Adams, B., Holmes, E.C., Zhang, C., Mammen Jr., M.P., Nimmannitya, S., Kalayanar-ooj, S., Boots, M., 2006. Cross-protective immunity can account for the alternating epidemic pattern of dengue virus serotypes circulating in Bangkok. *Proc. Natl. Acad. Sci. U.S.A.* 103 (38), 14234–14239.
- Aquino, V.H., Amarilla, A.A., Alfonso, H.L., Batista, W.C., Figueiredo, L.T., 2009. New genotype of dengue type 3 virus circulating in Brazil and Colombia showed a close relationship to old Asian viruses. *PLoS One* 4 (10), e7299.
- Araújo, J.M.G., Nogueira, R.M.R., Schatzmayr, H.G., Zanotto, P.M.A., Bello, G., 2008. Phylogeography and evolutionary history of dengue virus type 3. *Infect. Genet. Evol.* 9, 716–725.
- Barcelos Figueiredo, L., Batista Cecilio, A., Portela Ferreira, G., Paiva Drumond, B., Germano de Oliveira, J., Bonjardim, C.A., Peregrino Ferreira, P.C., Kroon, E.G., 2008. Dengue virus 3 genotype 1 associated with dengue fever and dengue hemorrhagic fever, Brazil. *Emerg. Infect. Dis.* 14, 314–316.
- Bennett, S.N., Holmes, E.C., Chirivella, M., Rodriguez, D.M., Beltran, M., Vorndam, V., Gubler, D.J., McMillan, W.O., 2003. Selection-driven evolution of emergent dengue virus. *Mol. Biol. Evol.* 20 (10), 1650–1658.
- Bennett, S.N., Drummond, A.J., Kapan, D.D., Suchard, M.A., Muñoz-Jordán, J.L., Pybus, O.G., Holmes, E.C., Gubler, D.J., 2010. Epidemic dynamics revealed in dengue evolution. *Mol. Biol. Evol.* 27 (4), 811–818.
- Carrington, C.V., Foster, J.E., Pybus, O.G., Bennett, S.N., Holmes, E.C., 2005. Invasion and maintenance of dengue virus type 2 and type 4 in the Americas. *J. Virol.* 79 (23), 14680–14687.
- Chatfield, C., 1975. *The Analysis of Time Series: Theory and Practice*, first ed. Chapman and Hall, London.
- Centers for Disease Control and Prevention, 1981. Dengue type 4 infections in U.S. travelers to the Caribbean. *MMWR Morb. Mortal. Wkly. Rep.* 30, 249–250.
- Cologna, R., Armstrong, P.M., Rico-Hesse, R., 2005. Selection for virulent dengue viruses occurs in humans and mosquitoes. *J. Virol.* 79 (2), 853–859.
- Dejnirattisai, W., Jumnainsong, A., Onsirakul, N., Fitton, P., Vasanawathana, S., Limpitikul, W., Puttikunt, C., Edwards, C., Duangchinda, T., Supasa, S., Chawansuntati, K., Malasit, P., Mongkolsapaya, J., Sreaton, G., 2010. Cross-reacting antibodies enhance dengue virus infection in humans. *Science* 328 (5979), 745–748.
- Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4, e88.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7, 214.
- Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. *Mol. Biol. Evol.* 22, 1185–1192.
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.
- Figueiredo, R.M., Naveca, F.G., Bastos, M.S., Melo, M.N., Viana, S.S., Mourão, M.P., Costa, C.A., Farias, I.P., 2008. Dengue virus type 4, Manaus Brazil. *Emerg. Infect. Dis.* 14 (4), 667–669.
- Foster, J.E., Bennett, S.N., Vaughan, H., Vorndam, V., McMillan, W.O., Carrington, C.V.F., 2003. Molecular evolution and phylogeny of dengue type 4 virus in the Caribbean. *Virology* 306, 126–134.
- Forshey, B.M., Morrison, A.C., Cruz, C., Rocha, C., Vilcarromero, S., Guevara, C., Camacho, D.E., Alava, A., Madrid, C., Beigolea, L., Suarez, V., Comach, G., Kochel, T.J., 2009. Dengue virus serotype 4, northeastern Peru, 2008. *Emerg. Infect. Dis.* 15 (11), 1815–1818.
- Gubler, D.J., 1998. Dengue and dengue hemorrhagic fever. *Clin. Microbiol. Rev.* 11 (3), 480–496.
- Guzman, M.G., Kouri, G.P., Bravo, J., Soler, M., Vazquez, S., Morier, L., 1990. Dengue hemorrhagic fever in Cuba 1981 a retrospective seroepidemiologic study. *Am. J. Trop. Med. Hyg.* 42, 179–184.
- Halstead, S.B., O'Rourke, E.J., 1977. Dengue viruses and mononuclear phagocytes. Infection enhancement by non-neutralizing antibody. *J. Exp. Med.* 146, 201.
- Halstead, S.B., Nimmannitya, S., Cohen, S.N., 1970. Observations related to pathogenesis of dengue hemorrhagic fever. IV. Relation of disease severity to antibody response and virus recovered. *Yale J. Biol. Med.* 42 (5), 311–328.
- Halstead, S.B., 2009. Antibodies determine virulence in dengue. *Ann. N.Y. Acad. Sci.* 1171 (Suppl. 1), E48–56.
- Hanna, J.N., Ritchie, S.A., Hills, S.L., Pyke, A.T., Montgomery, B.L., Richards, A.R., Piispanen, J.P., 2003. Dengue in north Queensland, 2002. *Commun. Dis. Intell.* 27 (3), 384–389.
- Henchal, E.A., Repik, P.M., McCown, J.M., Brandt, W.E., 1986. Identification of an antigenic and genetic variant of dengue-4 virus from the Caribbean. *Am. J. Trop. Med. Hyg.* 35 (2), 393–400.
- Holmes, E.C., 2004. The phylogeography of human viruses. *Mol. Ecol.* 13 (4), 745–756.
- Jansen, C.C., Beebe, N.W., 2010. The dengue vector *Aedes aegypti*: what comes next. *Microbes Infect.* 12 (4), 272–279.
- Kliks, S.C., Nimmannitya, S., Nisalak, A., Burke, D.S., 1988. Evidence that maternal dengue antibodies are important in the development of dengue hemorrhagic fever in infants. *Am. J. Trop. Med. Hyg.* 38 (2), 411–419.
- Klungthong, C., Zhang, C., Mammen Jr., M.P., Ubol, S., Holmes, E.C., 2004. The molecular epidemiology of dengue virus serotype 4 in Bangkok, Thailand. *Virology* 329 (1), 168–179.
- Lancioti, R.S., Gubler, D.J., Trent, D.W., 1997. Molecular evolution and phylogeny of dengue-4 viruses. *J. Gen. Virol.* 78 (9), 2279–2284.
- Lemey, P., Rambaut, A., Drummond, A.J., Suchard, M.A., 2009. Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* 5 (9), e1000520.
- Lourenço, J., Recker, M., 2010. Viral and Epidemiological determinants of the invasion dynamics of novel dengue genotypes. *PLoS Negl. Trop. Dis.* 4 (11), e894.
- Melo, F.L., Romano, C.M., Zanotto, P.M.A., 2009. Introduction of dengue virus 4 (DENV-4) genotype I into Brazil from Asia? *PLoS Negl. Trop. Dis.* 3 (4), e390.
- Messer, W.B., Gubler, D.J., Harris, E., Sivananthan, K., de Silva, A.M., 2003. Emergence and global spread of a dengue serotype 3, subtype III virus. *Emerg. Infect. Dis.* 9, 800–809.
- Pagel, M., Meade, A., Barker, D., 2004. Bayesian estimation of ancestral character states on phylogenies. *Syst. Biol.* 53, 673–684.
- Parker, J., Rambaut, A., Pybus, O.G., 2008. Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty. *Infect. Genet. Evol.* 8, 239–246.
- Pan American Health Organization, 1995–2010. Number of Reported Cases of Dengue and Dengue Hemorrhagic Fever (DHF), Region of the Americas (by country and subregion). Available from <http://www.paho.org/English/AD/DPC/CD/dengue.htm>.
- Posada, D., 2008. jModelTest: phylogenetic model averaging molecular biology and evolution. *Mol. Biol. Evol.* 27 (7), 1253–1256.
- Pybus, O.G., Rambaut, A., 2009. Evolutionary analysis of the dynamics of viral infectious disease. *Nat. Rev. Genet.* 10, 540–550.
- Rambaut, A., Drummond, A.J., 2007. Tracer v1.4, Available from <http://beast.bio.ed.ac.uk/Tracer>.
- Rico-Hesse, R., 1990. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology* 174, 479–493.
- Rodríguez, F., Oliver, J.L., Marín, A., Medina, J.R., 1990. The general stochastic model of nucleotide substitution. *J. Theor. Biol.* 142 (4), 485–501.
- Sangkawibha, N., Rojanasuphot, S., Ahandrik, S., Viriyapongse, S., Jatanasen, S., Salitul, V., Phanthumachinda, B., Halstead, S.B., 1984. Risk factors in dengue shock syndrome: a prospective epidemiologic study in Rayong, Thailand. I. The 1980 outbreak. *Am. J. Epidemiol.* 120 (5), 653–669.
- San Martín, J.L., Brathwaite, O., Zambrano, B., Solórzano, J.O., Bouckenoghe, A., Dayan, G.H., Guzmán, M.G., 2010. The epidemiology of dengue in the Americas over the last three decades: a worrisome reality. *Am. J. Trop. Med. Hyg.* 82 (1), 128–135.
- Thomas, L., Verlaeten, O., Cabié, A., Kaidomar, S., Moravie, V., Martial, J., Najioullah, F., Plumelle, Y., Fonteau, C., Dussart, P., Césaire, R., 2008. Influence of the dengue serotype, previous dengue infection, and plasma viral load on clinical presen-

- tation and outcome during a dengue-2 and dengue-4 co-epidemic. *Am. J. Trop. Med. Hyg.* 78 (6), 990–998.
- Twiddy, S.S., Holmes, E.C., Rambaut, A., 2003. Inferring the rate and time-scale of dengue virus evolution. *Mol. Biol. Evol.* 20 (1), 122–129.
- Ty Hang, V.T., Holmes, E.C., Duong, V., Nguyen, T.Q., Tran, T.H., Quail, M., Churcher, C., Parkhill, J., Cardoso, J., Farrar, J., Wills, B., Lennon, N.J., Birren, B.W., Buchy, P., Henn, M.R., Simmons, C.P., 2010. Emergence of the Asian 1 genotype of dengue virus serotype 2 in Viet Nam: in vivo fitness advantage and lineage replacement in South-East Asia. *PLoS Negl. Trop. Dis.* 20 (4/7), e757.
- Wang, E., Ni, H., Xu, R., Barret, A.D.T., Watowich, S.J., Gubler, D.J., Weaver, S.C., 2000. Evolutionary relationships of endemic/epidemic and sylvatic dengue viruses. *J. Virol.* 74 (7), 3227–3234.
- Weaver, S.C., Vasilakis, N., 2009. Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. *Infect. Genet. Evol.* 9, 523–540.
- Wilder-Smith, A., Gubler, D.J., 2008. Geographic expansion of dengue: the impact of international travel. *Med. Clin. North Am.* 92, 1377–1390.
- Zhang, C., Mammen Jr., M.P., Chinnawirotpisan, P., Klungthong, C., Rodpradit, P., Monkongdee, P., Nimmannitya, S., Kalayanaroj, S., Holmes, E.C., 2005. Clade replacements in dengue virus serotypes 1 and 3 are associated with changing serotype prevalence. *J. Virol.* 79 (24), 15123–15130.
- Zwickl, D.J., 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.

PHILOSOPHY OF BIOLOGY

Ending Microbial Myopia

Christian Julian Villabona-Arenas

A self-described “philosophical naturalist,” John Dupré aims “to draw—or at least support—philosophical conclusions from science” and to provide critical perspectives on it. In the 16 essays (published over the past decade) collected in *Processes of Life*, he applies his approach to contemporary biology. The volume stresses three points Dupré sees as crucial to understanding biology and evolution: the developmental systems perspective, epigenetics, and the importance of microbes. In its concluding section, he explores how the biological ideas he has discussed affect our conception of human nature.

Dupré (a philosopher of science at the University of Exeter) begins with a trio of chapters that survey his broader view of science, which he sees as produced by people in interaction with nature. He advocates a methodological pluralism and highlights the pitfalls of a purely reductionist enterprise. Given the variety of scientific activities, he wonders how they could be, in any sense, all about the same thing. Because methodologies have been developed in ways as diverse as the kinds of problems researchers pursue, the author outlines a set of epistemic virtues and advises practitioners to try to meet as many of them as possible.

One fundamental argument Dupré offers against reductionism is that biology works with concepts that depend not only on their constituents but also on the larger systems of which they are part. Debates over the nature of species have generated a substantial fraction of the evolutionary biology literature; less attention has been placed on the definition of genes. In addition to the concerns around such issues, there is the central question of what constitutes the individual organism. (With 90% of the cells within our bodies belonging to microbes, there is much to be said for considering that bacterial community as a single composite entity.) The author reminds us “that theoretical consider-

ations are insufficient to determine unequivocally the boundaries of biological objects. Sometimes, perhaps always, this must be done relative to a purpose.” Altogether, Dupré encourages trying to understand life as a hierarchy of dynamic and constantly changing processes. A living world where none of the entities that constitute an organism are static implies an interactive flux subtly different in every iteration but similar enough to be a distinctive process. Because biological concepts are static

abstractions from life processes and different abstractions provide different perspectives on these processes, we face considerable difficulties in reconciling satisfactory general concepts.

Although their scientific importance remains generally unappreciated, microbes

Dupré and coauthor Maureen A. O’Malley (University of Sydney) present a strong case for ending the myopia that leads us to undervalue microbes.

As Dupré and O’Malley note, paying proper attention to microbes has already yielded observations that may force us to rethink fundamental ideas about evolution. In light of the extent of lateral gene transfer, the “tree of life” seems better considered as a net. Studies of cooperation, development, competition, and communication among unicellular organisms have revealed that they can possess many of the characteristics used to define multicellularity. Given that the entities that form lineages are not always the same as those that form metabolic wholes, collaboration may be the central characteristic of living matter. Traditional organisms cannot be seen as “the” biological individuals on which selection operates.” Abandoning some theoretical commitments will open the way for further understanding nature.

The author’s reasons for recognizing the importance of the environment accentuate the relevance of epigenetics and developmental systems theory, areas that for a long time attracted little interest but have now become

Processes of Life
Essays in the Philosophy
of Biology
by John Dupré
Oxford University Press,
Oxford, 2012. 360 pp. \$55, £30.
ISBN 9780199691982.

Boundary-blurring biofilms. If bacteria in dental plaque (shown here in an SEM) “function as a coordinated, spatially organized and fully metabolically integrated microbial community” (1), Dupré wonders why we would “not consider this community, the organized functional whole, to constitute an organism.”

surpass macrobes (a term Dupré advocates for those organisms that are not bacteria, archaea, or protists) in their contribution to evolutionary history: the first three billion years of life on Earth were overwhelmingly microbial. Even today, they include most living things, exhibit a greater metabolic diversity, and inhabit a wider range of environments (including extremely harsh ones); their collaborative enterprise is so extensive that at the heart of every interface between multicellular eukaryotes and the external environment lies a complex multispecies microbial community. In three essays,

very active. Beyond doubt, Dupré emphasizes, the perpetuation of life from one generation to the next requires much more than simply the passage of DNA. He concludes that genomes do not merely store information. Because of their constant dynamic interaction with other constituents of the cell, their capacities depend not only on their sequence of base pairs. More important, those capacities are determined by the systems of which the DNA molecules are only part.

Turning his attention to humans, Dupré criticizes “Evolutionary Psychology.” He argues that evolution has had ample time

The reviewer is at the Laboratory of Molecular Evolution and Bioinformatics, Department of Microbiology, Biomedical Sciences Institute, University of São Paulo, Avenida Professor Lineu Prestes 1734, CEP: 05508-000, São Paulo, Brazil. E-mail: cj.villabona@usp.br

since the Stone Age to shift our behavior and that “the complexity of the developmental interactions between a wide variety of internal and external factors” (processes such as epigenetics, niche construction, and cultural evolution) means our genes cannot produce a “fixed and constant” human nature. He applies similar reasoning in rejecting claims

of genes for race, highlighting the mistake made in thinking of strings of DNA as having specific functions defined only in terms of phenotypic outcomes. For such reasons, he rejects genetic determinism.

Through their thorough arguments, the essays in *Processes of Life* challenge widely held assumptions about biology and evolu-

tion. Dupré provides a view of life grounded in recent research and current understanding. His perspective also reminds us how much we do not know.

References

1. P. D. Marsh, *Caries Res.* **38**, 204 (2004).

10.1126/science.1224870

SCIENCE AND CULTURE

#FramingTheArtofScience

Melissa A. C. Ingram

Gerald Weissmann's *Epigenetics in the Age of Twitter* is a compilation of essays that intertwine the profound connections of science and art in the context of our modern era, with its rapid proliferation of mobile apps and instantaneous digital chatter. Weissmann (New York University) uses epigenetics, short-term heritable changes in gene expression that are influenced by our environment, to illustrate the selective pressures in the real social network of Western civilization. An avid reader of *The FASEB Journal* may recognize portions of these chapters, as Weissmann has modified his editorials to create a cohesive and entertaining book. It describes current scientific and cultural trends, using little-known stories of obscure scientists as well as common mysteries (such as the correlation between stress and graying hair and whether early Coca-Cola contained cocaine) to illuminate the ongoing challenges scientists face in dealing with scrutiny and criticism, from colleagues and from our broader society.

The book begins with an introduction to the revolution of digital social media, with the growing presence of wikis and Twitter, the formatting of scientific papers, and how tradition is morphing as journals such as PLoS encourage comments and tweets. In a particularly entertaining section, Weissmann treats us to a fictitious set of reactions, including Rosalind Franklin's “YOU PEEKED,” inserted into Watson and Crick's iconic 1953 publication. Elsewhere, he presents other scientific communication issues, such as the relentless need to publish and the gender gap in authorship and editor positions.

Weissmann also does a fantastic job of highlighting the role of women in both the arts

and sciences. He offers excerpts about many prominent women, such as Hypatia, who in fifth-century Alexandria was the only female member of the academy, taught mathematics and philosophy, and was murdered by the mob for all she represented and political jealousy. Other examples—including Marie Curie, Carol Greider, and Elizabeth Blackburn, who have broken the Nobel Prize glass ceiling and inspired women all over the world—call attention to the accomplishments and struggles of women in the sciences.

Each essay provides its fair share of wit and satire, poignantly illustrating faults or

Epigenetics in the Age of Twitter

Pop Culture and Modern Science

by Gerald Weissmann

Bellevue Literary Press,
New York, 2012. 256 pp.

Paper, \$18.95.

ISBN 9781934137390.

curiosities in current scientific thought or public discourse of the scientific realm. Weissmann, an immunologist obviously well versed in vernacular, even pokes fun at his own field. In the essay “Inflammation Is Complicated: From Metchnikoff to Meryl Streep,” he satirizes the notoriously complex role of inflammation and self-destruction. “Our cells in combat are bathed by evil humors with acronyms worthy of the Pentagon—C5a, C3a, IFN, IL, MIF, etc.—to which they can react in both pro- and anti-inflammatory modes. Signals from ‘helper’ and ‘suppressor’ lymphocytes trigger cellular signals via JAKs and the STATs, the NF-kBs and so forth and so on. The CIA would indeed have trouble deciphering the battle orders.”

Interspersed within are subtle warnings that we ought not be so distracted by the speed of digital chats or the plethora of blogs that we fail to learn from the lessons of the past. Weissmann reminds both current and future scientists how easy it can be to make a difference: “a single observation, by a single scientist, can have a profound impact. It's a gift, so to speak, that's ours for the taking.” The author also cautions readers that while social media and pop culture may influence opinions of science, pop culture is not science—“the bottom line is that the world is round, humans evolved from an extinct species and Elvis is dead.”

This is the age of Twitter, where unrelated ideas, erratic thoughts, and momentary feelings come together to summarize a life story. Or in this collection of Weissmann's essays, how modern science is evolving in the face of ever-increasing social awareness and scrutiny.

10.1126/science.1225538

Electronic nerves. From Fritz Kahn's *Das Leben des Menschen* (Kosmos, Stuttgart, 1926).

The reviewer is at the Department of Genetics, Cell Biology, and Development, Institute for Translational Neuroscience, University of Minnesota, Minneapolis, MN 55414, USA. E-mail: corn0156@umn.edu

CREDIT: COURTESY BELLEVUE LITERARY PRESS

The burden of dengue: Jundiaí, Brazil – January 2010

CHRISTIAN JULIÁN VILLABONA ARENAS¹, ANA VITORIA BOTELHO², ANDREA CRISTINA BOTELHO³, SAULO DUARTE PASSOS⁴,
PAOLO MARINHO DE ANDRADE ZANOTTO⁵

¹Research Assistant, Laboratory of Molecular Evolution and Bioinformatics, Department of Microbiology, Biomedical Sciences Institute, Universidade de São Paulo, (USP), SP, Brazil

²Research Associate, Centro de Tecnologias Estratégicas do Nordeste, Recife, PE, Brazil

³Research Assistant, Laboratory of Pediatric Infectious Diseases, Department of Pediatrics, Faculdade de Medicina de Jundiaí (FMJ), Jundiaí, SP, Brazil

⁴Professor at Laboratory of Pediatric Infectious Diseases, Department of Pediatrics, FMJ, Jundiaí, SP, Brazil

⁵Professor at Laboratory of Molecular Evolution and Bioinformatics, Department of Microbiology, Biomedical Sciences Institute, USP, São Paulo, SP, Brazil

SUMMARY

Objective: To study the antibody prevalence against dengue in the municipality of Jundiaí, São Paulo, Brazil, due to the low number of official confirmed autochthonous cases.

Methods: A serological study on dengue infection was conducted during January 2010 and previous reports on dengue and entomological surveillance during that period were reviewed. **Results:** A prevalence of 7.8% IgG positive (68:876) was found. Furthermore, based on the detection of IgM antibodies in five samples, it was observed that the incidence of dengue in the city at the time of the survey contrasts with the absence of notifications by local health authorities over the same period of time. **Conclusion:** These results highlight the discrepancies between the actual and the detected number of dengue infections, possibly due to significant numbers of asymptomatic infections aggravated by difficulties with dengue clinical diagnosis.

Keywords: Dengue; epidemiology; serology.

©2012 Elsevier Editora Ltda. All rights reserved.

Study conducted at Laboratory
of Molecular Evolution and
Bioinformatics, Department of
Microbiology, Biomedical Sciences
Institute, Universidade de
São Paulo, São Paulo,
SP, Brazil

Submitted on: 09/29/2011
Approved on: 03/17/2012

Correspondence to:
Christian Julián Villabona Arenas
Av. Prof. Lineu Prestes, 1734
São Paulo – SP, Brazil
CEP: 05508-000
Phone: +55 (11) 3091-8453
Fax: +55 (11) 3091-7354
cj.villabona@usp.br

Conflict of interest: None.

INTRODUCTION

Dengue fever (DF) is caused by any of the four closely related flaviviruses transmitted by *Aedes aegypti* mosquitoes (serotypes 1 through 4)¹. Accordingly, the geographic distribution of *Aedes* mosquitoes globally matches the area of dengue transmission². Currently, the dengue virus (DENV) is the most common arbovirus causing a neglected human disease around the world. Furthermore, the number of cases has increased dramatically during the past three decades in the Americas³. Currently, there is no available dengue vaccine or antiviral therapy; therefore, dengue prevention is based almost entirely on vector control⁴.

The Pan American Health Organization (PAHO) reports that Brazil comprised approximately 46% of the reported cases in 2009 (528,883) and 60.65% of the cases in 2010 (1,004,392)⁵. Moreover, the State of São Paulo was among those that had the most reported cases, according to the 2010 dengue epidemiological report from the Brazilian Ministry of Health⁶. During that year, the State of São Paulo accumulated 188,593 cases, a sharp increase (2,096%) compared to 8,996 cases reported in 2009 (Centro de Vigilância Epidemiológica Alexandre Vranjac – CVE)⁷. Despite efforts to contain the disease, dengue keeps growing in most states of Brazil where the vector thrives.

In 2010, the municipality of Jundiaí, in the state of São Paulo, had an estimated population of 369,710⁸. Dengue cases have been documented since 2004 by the municipal health authorities, and there was a peak of 51 autochthonous cases in 2008. Interestingly, in 2009 only four autochthonous cases were reported.

The low number of confirmed autochthonous cases in Jundiaí suggests that there were cases that went undetected, either by difficulties with the diagnosis or because they were asymptomatic (in which case patients do not show up for testing), or as a result of broad herd immunity in the human population that does not allow for the onset of epidemics with many symptomatic cases. To address this issue, a serological survey following a random sampling was conducted during January 2010 (when virus activity is expected to occur in the Southern Hemisphere) and also reviewed previous technical reports on dengue and entomological surveillance from both the state of São Paulo and the city of Jundiaí.

POPULATION AND METHODS

Jundiaí is a well-developed city located in the eastern part of the State of São Paulo at 23° 11' 11" S and 46° 53' 03" W, at an elevation of 762 m above sea level, occupying a land area of 431.97 km² within at the Atlantic rainforest biome. According to the state of São Paulo's Metropolitan Planning Company (Empresa Paulista de Planejamento Metropolitano – Emplasa), Jundiaí is composed of an agglomeration of municipalities (Cabramatta, Limpo Paulista, Itupeva, Jarinu,

Jundiaí, Louveira, and Várzea Paulista) totaling 800,000 inhabitants, located between the important urban-industrial centers of São Paulo and Campinas, with about 18 million people within a radius of 80 km. Most of the dislocations that occur every day are directed towards the agglomeration itself, with the second destination being the metropolitan area of São Paulo. It is noteworthy that, in health grounds, all municipalities in the conurbation have a heavy transit flow with Jundiaí. Most treatments outside the municipality of residence that require hospitalizations occur in health facilities of Jundiaí in the following proportions: Várzea Paulista (91.7%), Jarinu (90.9%), Itupeva (86.8%), Limpo Paulista (86.2%), and Cabramatta (75.6%).

The approach to this dengue serological survey was to check the immunological status of the population of Novo Horizonte and São Camilo, which where the neighborhoods that reported most dengue cases during the 2008 outbreak (37 cases in Novo Horizonte and six cases in São Camilo). The sample size was calculated assuming a prevalence of 0.25% and the estimated population size of the municipality, which means that a dengue prevalence ~ 250 times higher than what can be inferred from technical reports issued by the health authorities in Jundiaí (~ 0.01%) was considered. Therefore, partitioned according to the estimated population size of each neighborhood, the experimental design entailed approximately 340 samples from São Camilo and 560 samples from Novo Horizonte, for a total expected sample size of 900. Households for sample collection were randomly selected from address records provided by the municipality. Health technicians randomly selected one resident per household, from which a single blood sample of 5 mL was collected. The ethical institutional review boards of Instituto de Ciências Biomédicas (ICB) - Universidade de São Paulo (USP) and Faculdade de Medicina de Jundiaí (FMJ) reviewed and approved the study. All participants volunteered to provide blood samples and age information for this study, after receiving a detailed explanation of the study and signing an informed consent.

The serological survey was carried out using the dengue IgG and IgM enzyme-linked immunosorbent assay (ELISA) (Bioeasy), following suppliers' recommendations. Data on entomological surveillance during 2009 were retrieved from the Superintendency of Endemic Diseases Control (Superintendência de Controle de Endemias – SUCEN). DENV technical reports were obtained from the CVE or provided by the health authorities of the municipality. A total of 886 samples were collected from January 11 to 25, 2010 (299 samples from São Camilo and 587 from Novo Horizonte). By comparing the sampling age distribution with that of the State of São Paulo⁹, it was found that both age structures were not significantly different ($\chi^2 = 0.0046$, d.f. = 7, $p = 1.0$), meaning that the sampling from Jundiaí could be considered a random sampling of the state population.

RESULTS AND DISCUSSION

ELISA was carried out for 876 serum samples, since ten samples from Novo Horizonte were not included due to technical problems. The 83 samples that were initially positive for IgG, were re-tested immediately after the first trial resulting in 68 confirmed positives (68:876) (Table 1). Subsequently, a serological test for dengue IgM antibodies was performed using the 68 IgG positive samples, resulting in five IgM seropositives (5:68). Remarkably, a 7.8% seropositivity for IgG in the two neighborhoods under study was found. Assuming that the sample reflects to some extent the local population, one could extrapolate a cumulative number of approximately 28,690 possible seropositive cases in Jundiaí. In contrast, the local authorities in Jundiaí reported a cumulative number of 98 autochthonous cases from 2004 to 2008. Likewise, a serological survey carried out in the region of the US-Mexican border revealed, for two cities, that the prevalence of the antibody against dengue was of 39% to 50%, a percentage 20 times higher than the cumulative number of cases from 1980 to 2007¹⁰. Discrepancies between reported dengue disease cases and actual level of transmission may arise due to silent transmission given the high asymptomatic infection rate¹¹. Another explanation foresees a scenario of re-introductions of viremic patients or vectors, which do not ignite large outbreaks. For example, five imported dengue cases were reported in Jundiaí during 2009, and the daily population commutes towards Jundiaí, indicating a strong interaction between the municipalities of the conurbation. In any case, it can be argued that proper verification of the actual number of infected people is paramount in localities with reduced number of notified cases, especially in countries such as Brazil where suspected dengue cases are among the diseases of compulsory notification. This notification does not require the confirmation of diagnosis, and all suspected cases do not have to be reported to and further investigated by the public health agencies. Notably, it has been shown that insufficient political commitment, inadequate financial resources, and increased urbanization have contributed to the dramatic increase in dengue cases in the Americas and, consequently, the epidemic control will largely be financially dictated, whereby budgetary allocations should match the size of the problem^{3,12}. Furthermore, an analysis combining available information on

reported cases, levels of underreporting, and cost per case, showed that Brazil is the country with highest economic impact of dengue, and that roughly 60% of these costs are due to productivity losses (indirect costs), which affect households, employers, and government¹³, highlighting a need for improved surveillance.

Given the disagreement between officially reported cases and the estimated and predicted seroprevalence, available statewide data on the yearly number of notified dengue cases and the yearly-averaged mosquito infestation indexes for Jundiaí and other 124 municipalities from 2009 was further analyzed. Previous studies^{14,15} were able to predict dengue transmission based on the Breteau Index (BI). The control of vector-borne diseases, intermediate hosts, and entomological surveillance performed by the SUCEN in the State of São Paulo is based on the BI and on a larval index (PI, which indicates the number of containers positive for larvae per 100 houses). Therefore, both BI and PI were contrasted with reported cases. Although the infestation indexes and number of cases showed seasonal increase in summer months, a simple linear regression did not show any significant correlation between the number of total dengue cases to either the annual averaged BI ($r^2 = 0.003$) or PI ($r^2 = 0.003$). In any case, it is important to highlight that yearly-averaged infestation indexes for Jundiaí (IP = 0.93, IB = 0.94) were similar to those measured for nearby cities that registered higher numbers of cases, such as Campinas (IP = 0.99, IB = 1.08, with 168 autochthonous and 22 imported cases). These data, in agreement with the present serology data, suggests that the real number of infections in Jundiaí (and maybe elsewhere) may be considerably higher than the detected by the health authorities. The numbers of dengue cases (autochthonous and imported) reported by the CVE, which varied quite significantly across the State, were also considered. As an example, in the 42 cities within the administrative area XVII (where Jundiaí belongs), the number of both types of cases had a strongly right-skewed distribution (skewness coefficient ≈ 4.7) with a median of zero cases for both imported (range = 22) and autochthonous (range = 168) cases, with mean values of 9.3 ± 4.3 , and 1.7 ± 0.6 respectively. Other than low numbers, these results evidenced a significant variation on the actual figures at nearby cities, which means that there is a great locality-specific fluctuation in reported cases for

Table 1 – Serological results for IgG and IgM antibodies during January 2010 in two neighborhoods of Jundiaí, São Paulo, Brazil

ELISA	São Camilo		Novo Horizonte		Total	
	Tested	Positive	Tested	Positive	Tested	Positive
IgG	299	20 (6.7%*)	577	48 (8.3%)	876	68 (7.8%)
IgM	20	1 (5%)	48	4 (8.3%)	68	5 (7.3%)

*Percentages of positives from the number of samples tested, respectively; ELISA, enzyme-linked immunosorbent assay.

neighboring cities with similar demographic and ecological characteristics to sustain dengue transmission.

However, from the present data alone, it cannot be concluded that dengue cases must have occurred and were underreported in Jundiaí, because the 68 seropositive individuals were not interviewed for a history of dengue-like illness in the past five to ten years. Consequently, it cannot be asserted that there is a sub-notification of cases. As an additional caveat, it is worth considering that the IgG negative samples were not tested for IgM. Therefore, a proper distinction between primary or secondary infections was not possible. Nevertheless, IgM antibodies are usually detected during the first three months post-infections. Additionally the present findings may also imply five recent infection events, which had to have occurred between November 2009 and January 2010 and could be significant for a sustained silent transmission of dengue.

Remarkably, the appropriate method to estimate sub-notification would be by reviewing medical records of acute febrile illnesses at the primary care facilities and identifying those patients with clinical symptoms compatible with acute dengue fever, then addressing why dengue was considered. Moreover, in- and out-patient records from Jundiaí hospitals were not analyzed for febrile syndromes observed during 2009-2010. Nevertheless, perhaps some dengue cases were diagnosed as influenza following the H1N1 flu pandemic. Given that dengue infection with any of the dengue serotypes may be asymptomatic in the majority of cases or may result in a wide spectrum of clinical symptoms, ranging from a flu-like fever to an incapacitating high fever, many dengue cases go undiagnosed. In addition, passive surveillance has been shown to significantly underestimate infection rates and is a factor constraining the control of dengue in the Americas¹².

ACKNOWLEDGEMENTS

CJVA holds a FAPESP scholarship (2011/17071-2), PMAZ holds a CNPq-PQ scholarship and AVB was a postdoctoral fellow (FAPESP scholarship 2008/01636-8) at the time of the study. The authors thank Teresa Maria Furlan for providing reagents and Antonio Henrique Alves Gomes for providing mosquito surveillance data. The authors are grateful to the staff of Medical Hospital of the Medical School of Jundiaí for their help with collecting samples and to the health authorities of Jundiaí for their lasting support and interest in this work.

REFERENCES

1. Gubler DJ. Epidemic dengue/dengue haemorrhagic fever: a global public health problem in the 21st century. *Dengue Bull.* 1997;21:1-15.
2. Urdaneta-Marquez L, Failloux AB. Population genetic structure of *Aedes aegypti*, the principal vector of dengue viruses. *Infect Genet Evol.* 2010;11(2):253-61.
3. San Martín JL, Brathwaite O, Zambrano B, Solórzano JO, Bouckennooghe A, Dayan GH, et al. The epidemiology of dengue in the Americas over the last three decades: a worrisome reality. *Am J Trop Med Hyg.* 2010;82(1):128-35.
4. Farrar J, Focks D, Gubler D, Barrera R, Guzman MG, Simmons C, et al. Towards a global dengue research agenda. *Trop Med Int Health.* 2007;12(6):695-9.
5. Pan American Health Organization. Number of reported cases of dengue and figures from 1995 to 2011 (to week noted by each country). [cited 2011 Aug 25]. Available from: <http://new.paho.org/hq>.
6. Organización Panamericana de la Salud. Actualización regional fengue: programa regional de dengue. [cited 4 Feb 2010]. Available from: <http://new.paho.org/hq/>.
7. Centro de Vigilância Epidemiológica "Alexandre Vranjac". Distribuição dos casos de dengue autóctones segundo o município provável de infecção e casos importados de outros estados segundo o município de residência no Estado de São Paulo no período de janeiro a dezembro de 2009-2010 [cited 25 Aug 2011]. Available from: http://www.cve.saude.sp.gov.br/hqm/zoo/den09_import_autoc.htm.
8. Fundação vinculada à Secretaria Estadual de Planejamento e Desenvolvimento Regional do Estado de São Paulo - SEADE, www.seade.gov.br.
9. Brazilian Institute of Geography and Statistics – Instituto Brasileiro de Geografia e Estatística – IBGE, www.ibge.gov.br.
10. Ramos MM, Mohammed H, Zielinski-Gutierrez E, Hayden MH, Robles López JL, Fournier M, et al. Epidemic dengue and dengue hemorrhagic fever at the Texas-Mexico border: results of a household-based seroepidemiologic survey, December 2005. *Am J Trop Med Hyg.* 2008;78(3):364-9.
11. Endy TP, Yoon I, Mammen M. Prospective cohort studies of dengue viral transmission and severity of disease. *Curr Top Microbiol Immunol.* 2010;338:1-13.
12. Tapia-Conyer R, Méndez-Galván JF, Gallardo-Rincón H. The growing burden of dengue in Latin America. *J Clin Virol.* 2009;46(S2):S3-S6.
13. Shepard DS, Coudeville L, Halasa YA, Zambrano B, Dayan GH. Economic impact of dengue illness in the Americas. *Am J Trop Med Hyg.* 2011;82(2):200-7.
14. Sanchez L, Cortinas J, Pelaez O, Gutierrez H, Concepción D, Van de Stuyt P. Breteau Index threshold levels indicating risk for dengue transmission in areas with low *Aedes* infestation. *Trop Med Int Health.* 2010;15(2):173-5.
15. Sanchez L, Vanlerberghe V, Alfonso L, Marquetti MDC, Guzman MG, Bisset J, et al. *Aedes aegypti* larval indices and risk for dengue epidemics. *Emerg Infect Dis.* 2006;12(5):800-6.

Worldwide Spread of Dengue Virus Type 1

Christian Julián Villabona-Arenas, Paolo Marinho de Andrade Zanotto*

Laboratory of Molecular Evolution and Bioinformatics, Department of Microbiology, Biomedical Sciences Institute, University of São Paulo, São Paulo, Brazil

Abstract

Background: DENV-1 is one of the four viral serotypes that causes Dengue, the most common mosquito-borne viral disease of humans. The prevalence of these viruses has grown in recent decades and is now present in more than 100 countries. Limited studies document the spread of DENV-1 over the world despite its importance for human health.

Methodology/Principal Findings: We used representative DENV-1 envelope gene sequences to unravel the dynamics of viral diffusion under a Bayesian phylogeographic approach. Data included strains from 45 distinct geographic locations isolated from 1944 to 2009. The estimated mean rate of nucleotide substitution was 6.56×10^{-4} substitutions/site/year. The larger genotypes (I, IV and V) had a distinctive phylogenetic structure and since 1990 they experienced effective population size oscillations. Thailand and Indonesia represented the main sources of strains for neighboring countries. Besides, Asia broadcast lineages into the Americas and the Pacific region that diverged in isolation. Also, a transmission network analysis revealed the pivotal role of Indochina in the global diffusion of DENV-1 and of the Caribbean in the diffusion over the Americas.

Conclusions/Significance: The study summarizes the spatiotemporal DENV-1 worldwide spread that may help disease control.

Citation: Villabona-Arenas CJ, Zanotto PMda (2013) Worldwide Spread of Dengue Virus Type 1. PLoS ONE 8(5): e62649. doi:10.1371/journal.pone.0062649

Editor: Nikos Vasilakis, University of Texas Medical Branch, United States of America

Received: January 30, 2013; **Accepted:** March 24, 2013; **Published:** May 13, 2013

Copyright: © 2013 Villabona-Arenas, Zanotto. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) grant 2010/19059-7. CJVA holds the FAPESP scholarship 2011/17071-2 and PMAZ holds a CNPq - PQ scholarship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: pzanotto@usp.br

Introduction

Dengue virus type 1 (DENV-1) is one of the four serotypes of the arthropod borne viruses (arbovirus) that causes Dengue [1,2]. Mosquitoes (mainly of the *Aedes* genus) transmit the virus and also infect non-human primates in a sylvatic cycle [3–5]. This RNA virus causes a fever that may be temporarily incapacitating and characterized by extreme pain and stiffness of the joints, hence known as break-bone fever. Other more severe manifestations of dengue infection, such as dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS) can be fatal if unrecognized and not properly treated in a timely manner. The prevalence of this disease has grown in recent decades and is now endemic in more than 100 countries [6].

The World Health Organization lists Dengue as a tropical neglected disease with other parasitic and bacterial infections [7,8]. The cumulative reported numbers of dengue cases in 2012 showcase the devastating impact of dengue on human health around the world: 262,542 cases (data up to week ending 14 Nov) in the western pacific region [9] and 992,285 cases (data up to epidemiological week 43) in the Americas [10]. Furthermore, the economic cost of dengue in the Americas reached US\$ 2.1 billion per year [11] and in Puerto Rico the aggregate annual cost averaged US\$ 38.7 million [12]. A clear picture of the dengue global geographic expansion is of consequence, because vector control remains as the main resource for disease containment, since the leading vaccine candidate was only 30 percent effective in its first large clinical trial [13].

RNA virus molecular evolution is fast and occurs at approximately the same temporal scale of the ecological processes that shape their diversity [14–16]. This sanctions the use of Bayesian methods to phylodynamics [17], an approach that infer the origins and distribution of different viral strains, which are particularly useful for healthcare agents that may prefer maps and trends rather than raw numbers to communicate health emergencies [18,19].

Despite high-endemicity of dengue viruses in tropical and subtropical countries, relatively few studies document viral dynamics and patterns of spread among human populations [20–25]. For DENV-1, Chen and Vasilakis [26] unrevealed biogeographical patterns of dispersal using 1812 strains based and taxa associations from the topology of tree. That study highlighted the role of Southeast Asian countries as sources of dengue epidemics and suggested a pattern of evolution radiating around spatially-defined geographic clades. Moreover, Allicock *et al.* [27] reconstructed the spatiotemporal spread of the virus in the Americas with 109 genotype V strains and documented the Lesser Antilles as the geographic location of ancestral infection. Therefore, in order to expand the knowledge on DENV-1 molecular epidemiology, by moving beyond the single place/time paradigm phylogenetic approach, bringing together a large number of sequences and providing a statistic comprehensive description on the genetic movement and distribution, the aim of this study was to reconstruct the temporal and spatial phylody-

namics of different genotypes of DENV-1 and to stress its importance for control.

Materials and Methods

Sequence Data

Complete envelope sequences (E gene) of DENV-1 with known time (year) and geographical locality (country) of isolation were retrieved from GenBank (1583 envelope gene sequences from 45 distinct geographic locations, data up to Jan 2009). A gene approach was preferred over a genomic one because of high computational demands posed by model-based algorithms when handling larger sequences and also because much more sequences are available for that gene, especially for the old strains, providing a less biased dataset. However, the envelope is a major determinant of tropism and the primary target of virus-neutralizing antibodies and several previous studies have demonstrated its contribution to DENV phylogenetics [28]. Recombinant clones were removed for our analysis and pairwise homoplasy index (PHI) calculations did not show evidence for recombination in the final dataset ($p=0.2224$) [29]. Sequences were aligned using Muscle software [30,31] followed by visual inspection and manual editing with the Se-Al v2.0 program (<http://tree.bio.ed.ac.uk/software/seal/>).

Genotyping

In order to classify taxa and to create datasets for each genotype a Maximum likelihood (ML) tree was inferred using Garli v2.0 software and the 1583 E sequences [32]; bootstrapping values for branches were estimated with 100 non-parametric replicates. This resulted in 1189 strains from Genotype I, one from Genotype II, four from Genotype III, 81 from Genotype IV and 306 from genotype V (Figure 1). Due to the large number of sequences available for Genotype I (773 being isolates from Vietnam), monophyletic groups with sequences from a single country were collapsed to include only the oldest taxon, the most recent and one halfway taxa (If there were several sequences at each of these time point, it was randomly selected). This procedure allowed the creation of a genotype I down sampled data set (180 taxa) that was used hereafter and was minimally representative of all possible migratory events over time. Also, a down sampled dataset for DENV-1 was built to estimate evolutionary parameters for the serotype (this exploration using over 500 sequences was computationally impractical in a Bayesian framework) by selecting 30 sequences representative in time and space from genotypes I, IV and V and all the available sequences from genotypes II and III.

Likelihood Mapping

The phylogenetic signal of all datasets (full and down-sampled) was investigated by means of the likelihood mapping analysis of 10 million random quartets using TreePuzzle software [33]. For each quartet the three possible unrooted trees were reconstructed following a ML approach under the GTR+G+I substitution model, selected under the Akaike information criterion with JModeltest software [34]. The results were plotted as a triangle with fully resolved quartets shown as dots near the vertices and the unresolved quartets in the center of the triangle. Consequently, the relative phylogenetic signal in a map is directly proportional to the percentage of the concentration of dots at the vertices.

Bayesian Phylodynamics Reconstruction

The usefulness and justifications for studying dengue systematics and evolution using Bayesian-based methods are argued for elsewhere [35–39]. Tree topologies, model parameters, evolutionary rate and viral population size variation over time were co-estimated independently for the down-sampled DENV-1 (95 taxa) and genotypes I (180 taxa), IV (81 taxa) and V(306 taxa) datasets with an uncorrelated log-normal relaxed clock model and the MCMC method implemented in the Beast package version 1.7 [35]. Genotypes II and III were not included because they both had less than five available sequences. A Bayesian skyline plot (an approach with no dependence on a pre-specified parametric model of demographic history) was used as a coalescent prior during the estimation over time of the change in $N_e g$ (the product of the effective population size and the generation length in years) [36]. The $N_e g$ parameter is a surrogate for the variation of viral population size and approximates the number of new infections. Importantly, $N_e g$ estimated from viral genealogies was shown to be proportional to the number of dengue cases in real epidemics [23]. The MCMC analysis was run four times and until convergence (with a minimum of 24 million generations sampling every 8000). We measured MCMC convergence by estimating the effective sampling size (ESS), after a 20% burn-in, using Tracer software version 1.5 (<http://tree.bio.ed.ac.uk/software/tracer/>). We estimated uncertainty as 95% high probability densities (95% HPD). All parameters showed ESS values >100. The obtained topologies were summarized in a maximum clade credibility (MCC) tree.

Bayesian Phylogeographic Analyses

The Beast program allows for an ancestral reconstruction of discrete states where the spatial diffusion of the time scaled genealogy is modeled as a continuous-time Markov chain (CTMC) process over discrete sampling locations [37]. We used the Bayesian stochastic search variable selection (BSSVS) approach, which assumes exchange rates in the CTMC to be zero with some prior probability, to find the most parsimonious set of rates explaining the diffusion process along the phylogenies for each genotype. We use the software Spread v1.0.4 [38] to visualize diffusion rates over time: each location was represented by two capital letters using the ISO 3166-1 alpha-code and coordinates corresponded to the centroids of each country; a Bayes factor (BF) test was run to identify the rates contributing to the migration path [37] and all rates yielded a substantial BF (>3), and rates with the higher values (>20) were listed in Table S2. We also used the BayesTraits software package (available at <http://www.evolution.rdg.ac.uk/>) to reconstruct, under a maximum likelihood approach, the viral migration rates among major geographical areas (coded as discrete multistate traits: Africa/Arabian Peninsula, Central America, East Asia, Indian Ocean, Indochina, Maritime Southeast Asia, Micronesia/Polynesia, South America and The Caribbean). These analyses were done reconstructing character state changes onto representative sample of trees collected during the stationary phase of the MCMC runs [39]. The inferred rates of change among character states were taken as surrogates for migration rates among main areas (nodes) and incorporated as attributes in a network analysis. The importance of each area for DENV-1 global spread was evaluated by estimating the betweenness centrality (BC) using the Cytoscape 2.8.3 program [40]. BC was chosen because it informs on the number of shortest paths from all vertices to all others that pass through a node in a network. Therefore, it is a more informative measure of global topologic importance of a node than connectivity (which indicates

Figure 1. Evolutionary history of DENV-1. The phylogeny was inferred using a Maximum likelihood analysis of 1583 sequences. Because of the large number of isolates branch tips were removed. Numbers correspond to bootstrapping values.
doi:10.1371/journal.pone.0062649.g001

the number of local connections of a node irrespective of its neighbor's connectivity properties) [41].

Results and Discussion

Evolutionary History of DENV1

We retrieved DENV-1 envelope sequences from GenBank (Table S1 lists the accession numbers) and estimated the phylogenetic noise in each data set. We assumed that all of them

contained phylogenetic signal (Figure S1) since none of them had more than 30% of unresolved quartets (i.e., noise), further supporting the adequacy of the phylogenetic inference methods that followed.

The Full DENV-1 dataset used in this study came from 45 geographic distinct locations, ranging from Hawaii (EU848545 being a pooled serum from six patients in 1944) to a single 2009 isolate from Cambodia. The ML tree (Figure 1) showed the five genotypes previously proposed [28,42]: (i) Genotype I included

strains from Southeast Asia and East Africa, (ii) Genotype II included a strain from Thailand collected in 1963, (iii) Genotype III included the sylvatic strain collected in Malaysia and the strain FN825674 isolated in 2005, (iv) Genotype IV included strains from Southeast Asia, the West Pacific islands, Australia and a few from the Americas, (v) Genotype V represented all strains collected in the Americas, West Africa, and a limited number of strains collected from Asia. The two closely-related Hawaiian strains isolated in 1944 and 1945 had a 5.7% estimated sequence divergence when compared to Genotype I. Hence, we grouped them together following the canonical definition of a genotype: a monophyletic group of dengue viruses having no more than 6% sequence divergence among them (Table S3) [43].

The larger genotype datasets were used for phylogenetic reconstruction. Figure 2 shows the summarized the MCC trees for each genotype. There was no evident monophyly by country (Except for an Arabian clade within Genotype I) or by year of isolation and most clades had posterior probabilities values of 1.0. We argue that, even when studies showed monophyletic grouping by time and place for other serotypes [22,44], in our case, the inclusion of more sequences within a spatiotemporal context revealed greater distribution and circulation of lineages among localities. For instance, well-sampled locations appeared throughout some topologies: Thailand appeared all over the genotype I tree, whilst strains from Indonesia appeared all over the genotype IV tree. This justifies the use of large samples to satisfactorily address viral diffusion.

Rate of Nucleotide Substitution and TMRCA

The mean rate of nucleotide substitution was estimated using 95 taxa (30 sequences from Genotype I, IV and V; one sequence from genotype II and four sequences from genotype III) and resulted in 6.56×10^{-4} substitutions/site/year (Table 1). Rate values were compatible with those reported for dengue ($6-8 \times 10^{-4}$ substitutions/site/year [28]) and related flaviviruses [45–47].

DENV-1 sequences appeared to evolve under a non-strict albeit clock-like manner because the estimated coefficient of variation under uncorrelated substitution rates was 0.381 (A value higher than zero supports the assumption of rate heterogeneity). A comparable value was obtained for DENV-4 [48]. Therefore, a relaxed clock model was used in the calculation of rates of substitution for each genotype separately. The estimated time to the most recent common ancestor (TMRCA) was 109 years (95% HPD 91.20–129.75) (Table 1). Remarkably, the TMRCA for each genotype was similar (~ 72.5 years).

DENV-1 Phylogeography

We used a discrete phylogeography analysis to provide a somewhat clear picture of its phylogeography: the interactive KML files (Dataset S1, S2 and S3) display the inferred distribution history over time for each genotype and Figure 3 presents a summary of the plausible migratory paths. An assumption of the discrete model of location change is that ancestral viruses necessarily reside at only the sampled locations of the extant viruses, hence, the methods describe the underlying spatial dynamics more accurately as the sampling density increase [37]. The DENV-1 sequence data studied, including the down-sampled genotype I, were diverse across a larger set of geographic locations. However, we are aware that most countries experiencing dengue epidemics do not have published sequences nor the sampling is homogeneous over time. Therefore, given the limitations and the uncertainty associated with each state reconstruction, in some cases instead of pointing to a specific location, we referred to the region to which it belongs.

Distribution of Genotype I in Asia

Genotype I was restricted to Asia with an incursion into the Arabian Peninsula (KML S1 and Figures 2A and 3A; full MCC tree file available upon request). Taxa from the same country and year grouped in different clades. For example, we identified at least four different paraphyletic Cambodian clades associated with lineages from Vietnam. Duong *et al.* [49] also reported distinct Cambodian lineages and Raghvani *et al.* [50] reported that Vietnam received Cambodian lineages on different times. Altogether, these findings evidence a scenario where multiple lineages circulated in the same spatiotemporal context.

Hawaii appeared as the source of genotype I (state probability 0.83) but we argue that this result is biased because Hawaiian sequences are the oldest available so far: before and during World War II, dengue frequently occurred in US military personnel in Asia and the South Pacific, and dengue introduction into Honolulu occurred 1943 due to infected commercial airline pilots from the South Pacific [51]. Furthermore, the DENV-1 reference strain Mochizuki was isolated in 1943 in Japan [52]. The dominant location throughout the phylogeny and hence the more likely hub of diffusion is Thailand and together with its neighbors represent the most likely genotype I source. Actually, over the past decades, the number of cases increased steadily in that region [53].

Distribution of Genotype IV in Asia and the Pacific

Genotype IV was found in Asia and the Pacific with an incursion into the Indian Ocean and the Americas (KML S2 and Figures 2B and 3B); Australian sequences also clustered within this genotype. Our findings suggest that the Philippines were probably the ancestral area for this genotype (state probability 0.14, other states with probabilities < 0.07) and the source for both, one of the introductions into the Pacific and for the incursion into the Americas. Even so, the maritime country of Indonesia, the dominant location throughout the phylogeny, was the likely source for the spreading all over the region and into the Indian Ocean countries.

The Indian Ocean and Polynesians clades suggested that lineages diverged considerably for sufficient time after introduction, perhaps due to its geographical distance. Furthermore, their short branch lengths suggest rapid dissemination of viruses, possibly facilitated by vector competence [54] and host population susceptibility. We also found that strains from French Polynesia were associated with both, genotypes IV and V (Figure 2C). These findings are in line with phylogenetic studies that suggested multiple introductions of viruses from different genotypes into the Pacific [55].

When comparing Asian genotypes (Genotypes I and IV) it seems that genotype I distributes mostly over Indochina, while genotype IV distributes mostly over the Maritime Southeast Asia and Pacific islands, such as the Federated States of Micronesia (FM). Remarkably, some countries have reported different genotypes during the same outbreak (*e.g.*, Indonesia, Singapore and Vietnam), which begs the question of a possible role of intra-genotype competition impacting on genetic DENV diversity as suggested for Asian DENV-4 genotypes [24]. Further studies are crucial because about 400,000 cases of severe dengue, a leading cause of childhood mortality, are reported yearly in Southeast Asia with Thailand and Indonesia leading in number of cases [56].

While dengue is more prevalent in Southeast Asia and the Western Pacific region, other Asian countries also experienced DENV-1 outbreaks. For example, Guangdong – China's most populated province – experienced outbreaks by DENV-1 since 1990 [57,58]. In our study, samples from China obtained from 1993 to 2002 were classified as genotype IV, while the oldest

A

B

C

Figure 2. The Phylogenetic relationships of DENV-1 genotypes. Topologies correspond to Maximum Clade Credibility trees (MCC) for A) Genotype I, B) Genotype IV and C) Genotype V. Because of the large number of isolates branch tips were removed. Major geographical clades are indicated (ISO 3166 Codes). Black numbers correspond to posterior probability estimates and gray numbers correspond to the TMRCA. Scale bar for years. Triangles represent Thailand isolates and squares represent Indonesia isolates.
doi:10.1371/journal.pone.0062649.g002

Chinese sequence from 1980 and the most recent ones from 1998 to 2006 were classified as genotype I. Our results, as those of Chen [57], also suggested that Singapore, Thailand and Vietnam were plausible sources of DENV-1 for China, Cambodia and Malaysia. Moreover, studies with symptomatic travelers revealed that dengue DENV-1 strains are frequently introduced into Japan and Korea (90% of cases) from Southeast Asia [59,60] where the disease has been an important health problem since international travel became popular in developing countries.

Distribution of Genotype V in the Americas and the Caribbean

Genotype V comprises most American and Caribbean isolates and its wide distribution suggests *in situ* evolution among a susceptible human population after its introduction from Asia (KML S3 and Figures 2C and 3C). Clades belonged mostly to a specific region (Caribbean, South America or Central America). Carvalho *et al.* [56] showed the existence of various country-specific DENV-1 founder effects, which may have fostered geographical structuring in the Americas. We argue that the lack of a larger number of samples from most countries might be masking a greater exchange of virus as the ones displayed by Puerto Rico, Venezuela and Brazil.

We inferred the introduction of genotype V into the Caribbean around 1971 and of genotype IV into Central America around 1979 (in line with Allicock *et al.* [27] estimates). Extensive epidemic outbreaks of DENV-1 were documented in 1977 in Puerto Rico, northeastern Mexico and the United States (Texas) [61,62]. The low representation of Genotype IV over the region might be explained by cross-protection of immunity granted by the earlier exposure to Genotype V.

Our inferences showed that a strain from Southeast Asia (represented by Singapore with state probability 0.14 followed by Thailand with state probability 0.12) was related to the lineage introduced into the Caribbean (represented by Jamaica with state probability 0.22 followed by Grenade and Trinidad and Tobago with states probabilities 0.21 and 0.19 respectively) and therefore reinforces the notion that this region was an epicenter of the disease. The virus then moved into Grenada and Mexico and eventually to neighbor countries. We documented a second introduction, also from Singapore (state probability 0.23), into the British Virgin Islands and from there (state probability 0.76) to Brazil. The Pan American Health Organization attributed 20% of the Dengue cases in the Americas to DENV-1 during 1995–2007

and together with DENV-2 were the most frequent serotypes during the 1990s [63]. DENV-1 was detected for the first time in Brazil in 1981 and became a nationwide public health issue following the 1986 outbreak [64]. Outbreaks have been reported in Paraguay since 1988 and in Argentina from 1999 to 2000 in people traveling from Paraguay [65,66].

DENV-1 Distribution in Africa, Middle East and Europe

We documented DENV-1 dispersal into Africa, the Indian Ocean and the Arabian Peninsula. Different countries in Africa reported outbreaks, such as Djibouti, Côte d'Ivoire, Gabon, Kenya, Madagascar, Mayotte, Senegal, Sudan and Zanzibar [67,68]. Nonetheless, conditions in African countries preclude more frequent diagnostic testing and adequate surveillance, leading to dengue sub-notification that hampers a better understanding of the history and spread dynamics of DENV-1 in that continent. Critically, Africa's role in the worldwide spread of dengue cannot be underestimated. Domingo *et al.* [69] reported DENV-1 in samples from infected European travelers returning from dengue endemic areas from 2002 to 2008: Strains imported from Africa belonged to the three described genotypes, strains imported from Asia belonged to genotype I and IV and strains imported from Central America, the Caribbean and India belonged to genotype V.

A comprehensive Overview of the Worldwide DENV-1 Transmission Network

After obtaining a coarse picture of the spread of three main DENV-1 genotypes around the globe, we envisaged how the different world regions interconnect as a comprehensive web of transmission for the three main genotypes. Figure 4 depicts a spatially-adimensional, directed DENV-1 transmission network among major geographical areas. Node diameter and position are shown proportional to betweenness centrality (BC), which informs on the importance of the node in the whole transmission network. The arrows among nodes represent highly supported routes, while their thicknesses are proportional to the rates of viral movement among regions. This modeling exercise encapsulates the pivotal role of Indochina in the global epidemiology and transmission dynamics of DENV-1, as shown by its higher BC and the role of Maritime Southeast Asia in the spread of for genotype IV in the Americas. Moreover, East Asia (represented by southern China provinces) had extensive viral exchanges of Asian genotypes while in other localities, like Africa and the Arabian Peninsula, strains

Table 1. Estimated DENV-1 substitutions rates and the time to the most recent common ancestor (TMRCA) estimated under a Bayesian Skyline coalescent tree prior.

Set	Number of sequences	Rate (subs/site/year)	Coefficient of Variation	Rate 95% HPD	Node Age	95% HPD	TMRCA (Year)
DENV-1	95	6.56×10^{-4}	0.381	$5.32 \times 10^{-4} - 7.79 \times 10^{-4}$	109.17	91.20–129.75	1899
Genotype I	180	9.90×10^{-4}	0.408	$8.75 \times 10^{-4} - 1.11 \times 10^{-3}$	71.96	65.68–81.73	1937
Genotype IV	81	9.07×10^{-4}	0.906	$6.11 \times 10^{-4} - 1.26 \times 10^{-3}$	71.53	41.81–110.96	1936
Genotype V	306	6.71×10^{-4}	0.711	$5.70 \times 10^{-4} - 7.65 \times 10^{-4}$	73.41	56.73–93.20	1935

doi:10.1371/journal.pone.0062649.t001

Figure 3. The Spread of DENV-1 genotypes. The likely dispersion routes were estimated from a discrete diffusion phylogeography process for A) Genotype I, B) Genotype IV and C) Genotype V. Dotted lines represent plausible migratory paths. Gray acronyms represent the apparent source of origin for each genotype (ISO 3166 Codes). Gray lines represent inferred introductions by studies with diseased travelers. A location-annotated MCC to KML for visualization over time is available for Google Earth as supporting material.
doi:10.1371/journal.pone.0062649.g003

are apparently only introduced. Another interesting feature of this network is its modularity, since the Caribbean (second highest BC value after Indochina) did act as an important hub, mainly to Latin America while not exchanging significantly with Asia. Nevertheless, we must exercise caution when drawing conclusion about these findings because of the inferential limitations imposed by uneven spatial sampling in general.

Complex Demographic History of DENV-1 Genotypes

We reconstructed the population dynamics through time of the three main genotypes (I, IV and V). Figure 5 shows each demographic history indicated by superimposed skyline plots of *N_eg* that approximates the number of new infections over time. All plots had a maintenance phase until the 1970s when genetic diversity remained constant. The most conspicuous feature is the increase of oscillation in all skyline plots in the 1990's.

Genotype I *N_eg* exhibited two growing phases, one around 1994 and another one around 2002. Both phases were followed by stationary phases. Emerging lineages with epidemic potential, might explain this increase. For instance, Thai strains fell in multiples clades in the phylogeny (Figure 2A) and most of these fall into two different time periods: before 2001 and after 2002 coinciding with the second growing phase. Therefore, it seems that clade replacements occurred around those years and increased viral diversity. Crucially, a large outbreak occurred in Myanmar in 2001 [70] where DENV-1 displaced the other three serotypes.

Zhang *et al.* [71] documented DENV-1 clade replacements during the mid-1990s; events that coincided with the first growing phase. Studies showed that the replacements are associated with changes in serotype prevalence (a consequence of differential

susceptibility to cross-reactive immune responses) and enhanced mosquito transmission [71,72]. Recently, Yamanaka *et al.* [73] reported the displacement of Indonesian serotype 2, by DENV-1 Genotype I, with the proportion of severe cases increasing about three times.

Genotypes IV *N_eg* grew steadily since 1988 and experienced a severe bottleneck around 2000 that reflect a possible limited availability of susceptible hosts: DENV-2 have was the prevalent serotype in the largest cities of Indonesia from 1998 to 2008 [73,74] and also in the Philippines from 1995 to 2001 [75]. In addition, Asian DENV-4 increased incidence from 1996 to 2000 [24]. Nonetheless, the large outbreak over the Pacific during 2001 and 2002 that affected Palau, French Polynesia, Samoa, American Samoa, New Caledonia, Cook Islands, Tokelau, Wallis and Fortuna, Fiji and Solomon Islands and also reached Hawaii, Easter Island and Australia [76–78] might helped to overcome the bottleneck because of the number of new infections.

Genotype V *N_eg* started to grew steadily since 1973, reached a zenith during the late 1980s and decline after 1988. The higher *N_eg* values coincided with its ongoing distribution over the Americas and the Caribbean [79]. Moreover, the decline coincided with the introduction of DENV-3 genotype III into Panama in 1994, that spread throughout the Americas, displacing DENV-1 [80].

Revisiting the Origins of DENV-1

Ecologic and socio-demographic changes of the 20th Century favored the emergence of dengue, particularly the end of World War II led to uncontrolled urbanization where mosquitoes vector reached high densities and facilitated dispersal of dengue among

Figure 4. Network visualization of diffusion of DENV-1 genotypes. The analysis was done over major geographical areas and integrating migration rates among them as attributes. Thickness gauge at the bar scale for rates.
doi:10.1371/journal.pone.0062649.g004

diverse geographic regions, and of adequate conditions for the emergence of DHF in Southeast Asia [28,81]. Based on the oldest isolates, Mochizuki strain isolated in 1943 and the extant genotype II represented by Thailand isolated in 1963, we argue that DENV-1 evolved in Asia and later spread into Africa and the Americas [76]. Besides, Malaysia, is considered an area where all dengue serotypes evolved independently from a sylvatic ancestral lineage [82] and during a 2004 outbreak Teoh *et al.* documented a human DENV-1 strain that had >97% sequence similarity to the strain isolated from a sentinel monkey in 1972 [83]. Therefore, Southeast Asia still remains as an area sheltering a sylvatic mosquito transmission cycle.

Stoddard *et al.* [84] have shown that human movement is a critical, understudied behavioral component underlying the transmission dynamics of many vector-borne pathogens. Nonetheless, only a few studies address the issues of international travel that is responsible for the increasingly global widespread and complex distribution of the four DENV [50]. Wilder-Smith and Gubler [85] argued that epidemics, their seasonality, and oscillations over time are reflected by the epidemiology of dengue in travelers. Yet, the pattern of DENV-1 distribution presented here, seems to agree with that reported for DENV-4 in the Caribbean [86]. There, regional social and commercial activity relates to the exchange of viral strains, by means of the sporadic spread to

remote regions where, if unimpeded by host immunity, it will result in regional epidemics until host immunity acts as a barrier to continued regional viral flow. This scenario is worsened by the only brief protection against heterologous infection with a different serotype setting the endemic nature of dengue in the tropical regions by spatial-temporal travelling waves of the dengue disease [87,88].

At present, it seems likely that DENV activity will continue in tropical countries, as climate change, international travel, lack of vector control and urbanization, may all contribute to sustained epidemics and virus emergence and reemergence [89,90]. As an example, outbreaks of dengue occurred in Hawaii by an unknown serotype in the 1840s, by DENV-1 genotype I in 1943 and 1944 and by different strains of DENV-1 genotype IV in 2001 [91,92]. Once awareness's practices related to the disease have demonstrated adequate prevalence of preventive practices against the disease [93], the knowledge of the spatial-temporal dynamics of DENV-1 provides an important starting point for control initiatives by policy-makers. Understanding movement will facilitate identification of key localities for transmission which may provide targets for mosquito surveillance intensification and improved disease prevention.

In sum, we have compiled much of the available DENV-1 sequence data to unveil some of the history of their distribution

Figure 5. Demographic history for each genotype. Overlay of Bayesian skyline plot showing the $Ne.g.$ (product of the effective population size and the generation length) for each DENV-1 genotypes. Ninety-five percent high probability densities (95% HPD) are shown as thin discontinuous lines around each mean plot.

doi:10.1371/journal.pone.0062649.g005

around the world. We also reconstructed the demographic history of the genotypes and highlighted how spread and viral diversity make control of dengue a challenging task. Further studies are mandatory because dengue epidemics occur as a result of the emergence of new lineages and dissemination. Our results also highlight the need to overcome non-uniform sampling from distinct locations, especially from Africa. Despite the fact that dengue is regularly reported in several important regions, national surveillance data still significantly underestimate the true burden of the disease [94].

Supporting Information

Figure S1 Likelihood mapping diagrams for each genotype dataset. The relative phylogenetic signal is directly proportional to the percentage of the concentration of dots at the vertices.
(EPS)

Table S1 List of GenBank sequences used in this study.
(DOCX)

Table S2 The significant connections at Bayes factor 20 to establish epidemiological linkage in viral phylogeographic histories.
(DOCX)

Table S3 Estimates of Net Evolutionary Divergence between groups of sequences for the complete DENV-1 dataset. Analyses were conducted using the Maximum Com-

posite Likelihood model. The rate variation among sites was modeled with a gamma distribution (shape parameter = 0.294).
(DOCX)

Dataset S1 Interactive KML for Google Earth showing the reconstructed dispersal history over time for genotype I.
(RAR)

Dataset S2 Interactive KML for Google Earth showing the reconstructed dispersal history over time for genotype IV.
(RAR)

Dataset S3 Interactive KML for Google Earth showing the reconstructed dispersal history over time for genotype V.
(RAR)

Acknowledgments

We thank Caio Freire for assistance and writing PERL scripts for parallelizing the BayesTraits analyses and for two anonymous referees for their thorough reviews, criticisms and suggestions.

Author Contributions

Conceived and designed the experiments: CJVA PMAZ. Performed the experiments: CJVA. Analyzed the data: CJVA PMAZ. Contributed reagents/materials/analysis tools: CJVA PMAZ. Wrote the paper: CJVA PMAZ.

References

- Theiler M, Downs WG (1973) The arthropod-borne viruses of vertebrates. An account of The Rockefeller Foundation Virus Program, 1951–1970. London: Yale University Press. 606 p.
- Karabatsos N (1985) Supplement to International Catalogue of Arboviruses Including Certain Other Viruses of Vertebrates. San Antonio, Texas : American Society of Tropical Medicine and Hygiene.
- Fagbami AH, Monath TP, Fabiyi A (1977) Dengue virus infections in Nigeria: a survey for antibodies in monkeys and humans. *Trans R Soc Trop Med Hyg* 71: 60–65.
- Henchal EA, Putnak JR (1990) The dengue viruses. *Clin Microbiol Rev* 3: 376–396.
- Simmons CP, Farrar JJ, Nguyen VV, Wills B (2012) Dengue. *N Engl J Med* 366: 1423–1432.
- WHO (2013) Dengue. Available <http://www.who.int/topics/dengue>. Accessed 2013 Apr 1.
- WHO (2013) The 17 neglected tropical diseases. Available: http://www.who.int/neglected_diseases/diseases. Accessed 2013 Apr 1.
- CDC (2011) Neglected Tropical Diseases. Available <http://www.cdc.gov/globalhealth/ntd>. Accessed 2013 Apr 1.
- WPRO (2012) Dengue Situation Updates. Available: http://www.wpro.who.int/emerging_diseases/DengueSituationUpdates. Accessed 2013 Apr 1.
- PAHO (2012) Number of Reported Cases of Dengue and Figures for 2012 (to week noted by each country). Available <http://new.paho.org/hq>. Accessed 2013 Apr 1.
- Shepard DS, Coudeville L, Halasa YA, Zambrano B, Dayan GH (2011) Economic impact of dengue illness in the Americas. *Am J Trop Med Hyg* 84: 200–207.
- Halasa YA, Shepard DS, Zeng W (2012) Economic Cost of Dengue in Puerto Rico. *Am J Trop Med Hyg* 86: 745–752.
- Hirschler B (2012) Sanofi dengue vaccine less effective than hoped. Reuters. London: Reuters. Available: <http://www.reuters.com/resources/archive/us/20120911>. Accessed 2013 Apr 1.
- Zanotto PMA, Gould EA (2002) Molecular epidemiology, evolution and dispersal of the genus flavivirus. In: Leitner T, editor. *The molecular epidemiology of human viruses*. Norwell: Kluwer Academic Publishers. 167–195 p.
- Holmes EC (2004) The phylogeography of human viruses. *Mol Ecol* 13: 745–756.
- Holmes EC (2008) Evolutionary history and phylogeography of human viruses. *Annu Rev Microbiol* 62: 307–328.
- Ronquist F, Sanmartin I (2011) Phylogenetic Methods in Biogeography. *Annu Rev Ecol Syst* 42: 441–464.
- Duncombe J, Clements A, Hu W, Weinstein P, Ritchie S, et al. (2012) Geographical Information Systems for Dengue Surveillance. *Am J Trop Med Hyg* 86: 753–755.
- Fisher RP, Myers BA (2011) Free and simple GIS as appropriate for health mapping in a low resource setting: a case study in eastern Indonesia. *Int J Health Geogr* 10: 15.
- Zanotto P, Gould E, Gao G (1996) Population dynamics of flaviviruses revealed by molecular phylogenies. *Proc Natl Acad Sci U S A* 93: 548–553.
- Holmes E, Twiddy S (2003) The origin, emergence and evolutionary genetics of dengue virus. *Infect Genet Evol* 3: 19–28.
- Araújo JMG, Nogueira RMR, Schatzmayr HGH, Zanotto PMA, Bello G (2009) Phylogeography and evolutionary history of dengue virus type 3. *Infect Genet Evol* 9: 716–725.
- Mondini A, de Moraes Bronzoni RV, Nunes SHP, Chiaravalloti Neto F, Massad E, et al. (2009) Spatio-temporal tracking and phylodynamics of an urban dengue 3 outbreak in Sao Paulo, Brazil. *PLoS Negl Trop Dis* 3: e448.
- Villabona-Arenas CJ, Zanotto PMA (2011) Evolutionary history of Dengue virus type 4: Insights into genotype phylodynamics. *Infect Genet Evol* 11: 1–8.
- Twiddy SS, Pybus OG, Holmes EC (2003) Comparative population dynamics of mosquito-borne flaviviruses. *Infect Genet Evol* 3: 87–95.
- Chen R, Vasilakis N (2011) Dengue—quo tu et quo vadis? *Viruses* 3: 1562–1608.
- Allicock OM, Lemey P, Tatem AJ, Pybus OG, Bennett SN, et al. (2012) Phylogeography and Population Dynamics of Dengue Viruses in the Americas. *Mol Biol Evol* 29: 1533–1543.
- Weaver SC, Vasilakis N (2009) Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. *Infect Genet Evol* 9: 523–540.
- Bruen TC, Philippe H, Bryant D (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics* 172: 2665–2681.
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792–1797.
- Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics* 5: 113.
- Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Austin: The University of Texas.
- Strimmer K, Haesler AV (1997) Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Proc Natl Acad Sci U S A* 94: 6815–6819.
- Posada D (2008) jModelTest: phylogenetic model averaging. 25: 1253–1256.
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29: 1969–1973.
- Drummond A, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol* 22: 1185–1192.
- Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009) Bayesian phylogeography finds its roots. *PLoS Comput Biol* 5: e1000520.
- Bielejec F, Rambaut A, Suchard MA, Lemey P (2011) SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* 27: 2910–2912.
- Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Syst Biol* 53: 673–684.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13: 2498–2504.
- Smoot ME, Ono K, Ruschinski J, Wang P-L, Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27: 431–432.
- Gonzalez A, Escalante AA, Pujol FH, Ludert JE, Tovar D, et al. (2002) Diversity and Evolution of the Envelope Gene of Dengue Virus Type 1. *Virology* 303: 110–119.
- Rico-Hesse R (1990) Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology* 174: 479–493.
- Foster JE, Bennett SN, Carrington CVF, Vaughan H, McMillan WO (2004) Phylogeography and molecular evolution of dengue 2 in the Caribbean basin, 1981–2000. *Virology* 324: 48–59.
- Baillie G, Kolokotronis S, Waltari E (2008) Phylogenetic and evolutionary analyses of St. Louis encephalitis virus genomes. *Mol Phylogenet Evol* 47: 717–728.
- Jenkins GM, Rambaut A, Pybus OG, Holmes EC (2002) Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J Mol Evol* 54: 156–165.
- May FJ, Davis CT, Tesh RB, Barrett ADT (2011) Phylogeography of West Nile virus: from the cradle of evolution in Africa to Eurasia, Australia, and the Americas. *J Virol* 85: 2964–2974.
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol* 4: e88.
- Duong V, Simmons C, Gavotte L, Viari A, Ong S, et al. (2011) Genetic diversity and lineage dynamic of dengue virus serotype 1 (DENV-1) in Cambodia. *Infect Genet Evol* : journal of molecular epidemiology and evolutionary genetics in infectious diseases 1: 1–10.
- Raghwanji J, Rambaut A, Holmes EC, Hang VT, Hien TT, et al. (2011) Endemic dengue associated with the co-circulation of multiple viral lineages and localized density-dependent transmission. *PLoS Pathog* 7: e1002064.
- Gibbons RV, Streitz M, Babina T, Fried JR (2012) Dengue and US Military Operations from the Spanish – American War through Today. *Emerg Infect Dis* 18: 636–630.
- Hotta S (1952) Experimental studies on dengue I. Isolation, identification and modification of the virus. *J Infect Dis* 90: 1–9.
- Kantachuvessiri A (2002) Dengue hemorrhagic fever in Thai society. *Southeast Asian J Trop Med Public Health* 33: 56–62.
- A-Nuegonpipat A, Berlioz-Arthaud A, Chow V, Endy T, Lowry K, et al. (2004) Sustained transmission of dengue virus type 1 in the Pacific due to repeated introductions of different Asian strains. *Virology* 329: 505–512.
- Wu J-Y, Lun Z-R, James AA, Chen X-G (2010) Dengue Fever in mainland China. *The Am J Trop Med Hyg* 83: 664–671.
- Carvalho SES, Martin DP, Oliveira LM, Ribeiro BM, Nagata T (2010) Comparative analysis of American Dengue virus type 1 full-genome sequences. *Virus genes* 40: 60–66.
- Chen S (2011) The origin of dengue viruses caused the DF outbreak in Guangdong province, China, in 2006. *Infect Genet Evol* 11: 1183–1187.
- Mizuno Y, Kato Y, Kano S, Takasaki T (2012) Imported malaria and dengue fever in returned travelers in Japan from 2005 to 2010. *Travel Med Infect Dis* 10: 86–91.
- Ito M, Yamada K-I, Takasaki T, Pandey B, Nerome R, et al. (2007) Phylogenetic analysis of dengue viruses isolated from imported dengue patients: possible aid for determining the countries where infections occurred. *J Travel Med* 14: 233–244.
- SEARO (2012) Reported Cases of DF/DHF in Selected Countries in SEA Region (1985–2005). Available: http://209.61.208.233/en/Section10/Section332_1101.htm. Accessed 2013 Apr 1.
- PAHO (1981) Dengue in Mexico and in the United States, 1980. *Epidemiological Bulletin* 2: 9–11. Available http://www1.paho.org/english/dd/ais/EB_v2n1.pdf. Accessed 2013 Apr 1.
- Wilder-Smith A (2008) Geographic expansion of dengue: the impact of international travel. *Medical Clinics of North America* 92: 1377–1390.
- San Martín JL, Brathwaite O, Zambrano B, Solórzano JO, Bouckennooghe A, et al. (2010) The epidemiology of dengue in the Americas over the last three decades: a worrisome reality. *The Am J Trop Med Hyg* 82: 128–135.

64. Schatzmayr HG, Nogueira RMR, Travassos da Rosa APA (1986) An outbreak of dengue virus at Rio de Janeiro - 1986. *Mem Inst Oswaldo Cruz* 81: 245–246.
65. Avilés G, Rowe J, Meissner J, Manzur Caffarena JC, Enria D, et al. (2002) Phylogenetic relationships of dengue-1 viruses from Argentina and Paraguay. *Arch Virol* 147: 2075–2087.
66. Barrero PR, Mistchenko AS (2004) Complete genome sequencing of dengue virus type 1 isolated in Buenos Aires, Argentina. *Virus Res* 101: 135–145.
67. Amarasinghe A, Kuritsk JN, Letson GW, Margolis HS (2011) Dengue virus infection in Africa. *Emerg Infect Dis* 17: 1349–1354.
68. Laval Fd, Plumet S, Simon F (2012) Dengue Surveillance among French Military in Africa. *Emerg Infect Dis* 18: 342–343.
69. Domingo C, Niedrig M, Gascón J, Palacios G, Reyes N, et al. (2011) Molecular surveillance of circulating dengue genotypes through European travelers. *J Travel Med* 18: 183–190.
70. Thu H, Lowry K, Myint T (2004) Myanmar dengue outbreak associated with displacement of serotypes 2, 3, and 4 by dengue 1. *Emerg Infect Dis* 10: 593–597.
71. Zhang C, Mammen Jr MP, Chinnawirotpisan P, Klungthong C, Rodpradit P, et al. (2005) Clade replacements in dengue virus serotypes 1 and 3 are associated with changing serotype prevalence. *J Virol* 79: 15123.
72. Lambrechts L, Chevillon C, Albright RG, Thaisomboonsuk B, Richardson JH, et al. (2009) Genetic specificity and potential for local adaptation between dengue viruses and mosquito vectors. *BMC Evol Biol* 9: 160.
73. Yamanaka A, Mulyatno KC, Susilowati H, Hendrianto E, Ginting AP, et al. (2011) Displacement of the predominant dengue virus from type 2 to type 1 with a subsequent genotype shift from IV to I in Surabaya, Indonesia 2008–2010. *PLoS one* 6: e27322.
74. Setiati TE, Wagenaar JFP, Kruif MDD, Mairuhu ATA (2006) Changing Epidemiology of Dengue Haemorrhagic Fever in Indonesia. *Dengue Bulletin* 30: 1–14.
75. Salda LTD, Parquet MDC, Matias RR, Natividad FF, Kobayashi N, et al. (2005) Molecular epidemiology of dengue 2 viruses in the Philippines: genotype shift and local evolution. *The Am J Trop Med Hyg* 73: 796–802.
76. Cáceres C, Yung V, Araya P, Tognarelli J, Villagra E, et al. (2008) Complete nucleotide sequence analysis of a Dengue-1 virus isolated on Easter Island, Chile. *Arch Virol* 153: 1967–1970.
77. Lepers C (2002) DEN-1 Outbreak in the Pacific – Update as at 22 May 2002. Available: <http://www.spc.int/phs/ENGLISH/Publications/InformACTION/IA11/Dengue-update-E.pdf>. Accessed 2013 Apr 1.
78. Lepers C (2002) DEN-1 Outbreak in the Pacific– Update as of 17 January 2003 Heading: Melanesia. Available: <http://www.spc.int/phs/ENGLISH/Publications/InformACTION/IA13/DengueUpdate.pdf>. Accessed 2013 Apr 1.
79. Gubler D, Scheld W, Armstrong D (1998) Epidemic dengue and dengue hemorrhagic fever: a global public health problem in the 21st century. *Dengue Bulletin* 21: 1–14.
80. Messer WB, Gubler DJ, Harris E, Sivananthan K, de Silva AM (2003) Emergence and global spread of a dengue serotype 3, subtype III virus. *Emerg Infect Dis* 9: 800–809.
81. Hammon WM, Rudnick A, Sather GE (1960) Viruses Associated with Epidemic Hemorrhagic Fevers of the Philippines and Thailand. *Science* 131: 1102–1103.
82. Wang E, Ni H, Xu R, Barrett aD, Watowich SJ, et al. (2000) Evolutionary relationships of endemic/epidemic and sylvatic dengue viruses. *J Virol* 74: 3227–3234.
83. Teoh B, Sam S, Abd-Jamil J (2010) Isolation of ancestral sylvatic dengue virus type 1, Malaysia. *Emerg Infect Dis* 16: 1783–1785.
84. Stoddard ST, Morrison AC, Vazquez-Prokopec GM, Paz Soldan V, Kochel TJ, et al. (2009) The role of human movement in the transmission of vector-borne pathogens. *PLoS Negl Trop Dis* 3: e481.
85. Wilder-Smith A, Gubler DJ (2008) Geographic expansion of dengue: the impact of international travel. *Med Clin North Am* 92: 1377–1390.
86. Foster JE, Bennett SN, Vaughan H, Vorndam V, McMillan WO, et al. (2003) Molecular evolution and phylogeny of dengue type 4 virus in the Caribbean. *Virology* 306: 126–134.
87. Guzmán M, Kouri G, Bravo J (1990) Dengue hemorrhagic fever in Cuba, 1981: a retrospective seroepidemiologic study. *Am J Trop Med Hyg* 42: 179–184.
88. Sabin A (1952) Research on dengue during World War II. *The Am J Trop Med Hyg*: 30–50.
89. Leal EDS, Zanotto PM (2000) Viral diseases and human evolution. *Mem Inst Oswaldo Cruz* 95 Suppl 1: 193–200.
90. Jansen CC, Beebe NW (2010) The dengue vector *Aedes aegypti*: what comes next. *Microbes Infect* 12: 272–279.
91. Effler PV, Pang L, Kitsutani P, Vorndam V, Nakata M, et al. (2005) Dengue fever, Hawaii, 2001–2002. *Emerg Infect Dis* 11: 742–749.
92. Imrie a, Zhao Z, Bennett SN, Kitsutani P, Laille M, et al. (2006) Molecular epidemiology of dengue in the Pacific: introduction of two distinct strains of dengue virus type-1 [corrected] into Hawaii. *Ann Trop Med Parasitol* 100: 327–336.
93. Itrat A, Khan A, Javaid S, Kamal M, Khan H, et al. (2008) Knowledge, awareness and practices regarding dengue fever among the adult population of dengue hit cosmopolitan. *PLoS one* 3: e2620.
94. Wichmann O, Yoon I-K, Vong S, Limkittikul K, Gibbons RV, et al. (2011) Dengue in Thailand and Cambodia: an assessment of the degree of underrecognized disease burden based on reported cases. *PLoS Negl Trop Dis* 5: e996.

Genomic mosaicism in two strains of dengue virus type 3

Christian Julián Villabona-Arenas, Anderson Fernandes de Brito, Paolo Marinho de Andrade Zanotto*

Laboratory of Molecular Evolution and Bioinformatics, Department of Microbiology, Institute of Biomedical Sciences, University of São Paulo, Av. Prof. Lineu Prestes, 1734, CEP: 05508-000, São Paulo, SP, Brazil

ARTICLE INFO

Article history:

Received 5 March 2013

Received in revised form 14 May 2013

Accepted 15 May 2013

Available online 30 May 2013

Keywords:

Dengue virus type 3

Phylogeny

Recombination

Split decomposition

ABSTRACT

Recombination is a significant factor driving genomic evolution, but it is not well understood in Dengue virus. We used phylogenetic methods to search for recombination in 636 Dengue virus type 3 (DENV-3) genomes and unveiled complex recombination patterns in two strains, which appear to be the outcome of recombination between genotype II and genotype I parental DENV-3 lineages. Our findings of genomic mosaic structures suggest that strand switching during RNA synthesis may be involved in the generation of genetic diversity in dengue viruses.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The availability of increasing sequence information from different localities and refined phylogenetic tools make it possible to address many of the processes that lead to genetic diversification, adaptation and spread of *Dengue virus* (DENV) (Weaver and Vasilakis, 2009). Populations of RNA viruses usually harbor genetic diversity, associated with high mutation rates, small effective population and large population sizes (Simon-Loriere and Holmes, 2011). The exchange of genetic material is an important mechanism that allows adaptive change among natural populations and evidences its ubiquitous role in evolution (Awadalla, 2003).

Dengue fever in humans is caused by enveloped, positive single stranded RNA ([+] ssRNA) mosquito-borne viruses, which belong to the genus *Flavivirus* (Family: *Flaviviridae*) and comprise four serotypes (DENV-1 to DENV-4) (Karabatsos, 1985). Their genomes (~11 kbp) encode a single open reading frame, flanked by highly structured 5' and 3' untranslated regions (UTRs) (Lindenbach and Rice, 2003; Rice et al., 1985). The N-terminal of the polyprotein encodes the three structural proteins (C-prM-E), followed by seven non-structural (NS) proteins (NS1–NS2A–NS2B–NS3–NS4A–NS4B–NS5) (Chambers et al., 1990; Rice et al., 1985).

Concomitant infection with two different strains (a prerequisite for DENV recombination) was demonstrated both in humans and in mosquitoes (Bharaj et al., 2008; Loroño-Pino et al., 1999; Laille et al., 1991; Montenegro et al., 2006; Wang et al., 2003). Recently, Thai et al. (2012) documented distinct lineages within the same

host. Evidence of recombination between diverse dengue strains has arisen (Aaskov et al., 2007; AbuBakar et al., 2002; Carvalho et al., 2010; Chen et al., 2008; Craig et al., 2003; Lavergne et al., 2006; Uzcategui et al., 2001; Worobey et al., 1999) since the first report of recombination in natural populations in 1999 (Holmes et al., 1999). Also, recombination was shown in other [+] ssRNA viruses, (Becher and Tautz, 2011; Carney et al., 2011; Chuang and Chen, 2009; Gould et al., 2004; Kalinina et al., 2002; Reiter et al., 2011), Retroviruses (Nájera et al., 2002) and [–] ssRNA viruses (Sibold et al., 1999). Two main mechanisms have been proposed for viral RNA recombination: (i) non-replicative breakage and rejoining and (ii) replicative template switching (Cooper, 1974; Nagy and Simon, 1997). The first model is a viral replication-independent RNA recombination process in which RNAs spontaneously rearrange their sequences, as found in tissue culture cells infected by polioviruses (Gmyl et al., 1999) and pestiviruses (Becher and Tautz, 2011). The second mechanism predicts that during replication the viral RdRp pauses RNA synthesis and eventually jumps, together with the incomplete nascent RNA, to another template in order to continue RNA synthesis. This mechanism has also been studied for [+] RNA such as the *Brome mosaic virus* and polioviruses, pestiviruses, carmoviruses and tombusviruses (Arnold and Cameron, 1999; Cheng and Nagy, 2003; Dahourou et al., 2002; Duggal et al., 1997; Simon and Nagy, 1996).

Due to modern transportation, wide-range dispersion of viruses can take place within 24–36 h resulting in several epidemic outbreaks, some of which involve the co-circulation of different viral strains, which increases the likelihood of mixed infections (Gould and Gritsun, 2006). Herein, we made use of available data and distinct analytical approaches to identify for the first time recombinant viruses in whole-genome alignments of DENV-3.

* Corresponding author. Tel.: +55 11 3091 8453.

E-mail address: pzanotto@usp.br (P.M. de Andrade Zanotto).

2. Materials and methods

2.1. Data set

Genomic DENV-3 strains were retrieved from GenBank (Table S1, in the supplemental material) and their source feature inspected, as a caution required in obtaining data from a public database, to check each entry name and to exclude mutants and non-human/non-mosquitoes isolates. Alignments were built with the Muscle 3.8.31 program (Edgar, 2004a,b), manually inspected, and trimmed to include only coding regions.

2.2. Detecting recombination

We used the automated algorithms contained within the recombination detection program 3.0 (RDP3) (Martin et al., 2010) to screen for putative recombination events in the coding region alignment. We used the default settings for each algorithm with a highest acceptable p -value of 0.01 and the standard Bonferroni correction (a p -value modifier that decrease the p -value cutoff according to the size of the dataset being examined) and set the analysis to report all hits detected by 6 or more algorithms. Once a putative recombination event was detected, a smaller dataset was built from selected sequences that included the likely recombinant, surrogate parental taxa, and reference strains.

2.3. Detection of recombinant breakpoints in the whole-genome alignment

Two Hidden Markov Models procedures (HMM) were used to corroborate the results from the automated tests and accurately locate breakpoints: (i) The Husmeier and Wright approach (Husmeier and McGuire, 2003) with four taxa (a likely recombinant, the surrogate parental taxa and one reference strain) and a time-reversible F84 substitution model (Felsenstein and Churchill, 1995) implemented in the Topali 2.5 software (Milne et al., 2009); and (ii) the Westesson and Holmes approach (86) that follows the structural EM heuristic (Friedman et al., 2002) using the RecHMM software by setting the minimum cutoff length for a recombinant region to 400 bp and using two additional reference strains. Significant tree topology changes allow partition of the alignment into genomic regions with apparent distinct evolutionary histories. For instance, for four taxa, there are three different unrooted tree topologies; in the presence of recombination, the tree topology can change and thus become a random variable that depends on the sites under consideration (the hidden states of the HMM represent the three topologies).

2.4. Phylogenetic network analyses

Alignments were built for each detected recombinant, including the most plausible surrogate parental taxa and three reference sequences from each genotype (free of recombination signal as evidenced by further automated detection tests). A pairwise homoplasy index (PHI) based on the principle of refined incompatibility was calculated for each alignment to test for recombination (Bruen et al., 2006). Afterwards, Split decomposition analyses (Bandelt and Dress, 1992) were done to assess the presence of a 'phylogenetic network' as an indication of reticulate evolution. This approach enhances phylogenetic analyses of distance data by unveiling incompatible groupings of taxa (splits) that arise when recombination and/or horizontal gene transfer have occurred. Analyses were done with the SplitsTree 4.0 software (Huson and Klopper, 2005), with a bootstrapping of 100 replicates, using a general time reversible substitution

model (GTR). The BEAST software package (Drummond et al., 2012) was used to estimate the proportion of invariable sites (I) among site rate variation (γ); nucleotide frequencies; and substitution rates for the model.

2.5. Phylogenetic reconstruction

A Maximum Clade Credibility tree (MCC) and divergence times for each alignment were inferred by using a Bayesian Markov chain Monte Carlo (MCMC) approach as implemented in BEAST. The estimated GTR+I+ γ model was used together with a relaxed molecular clock (Drummond et al., 2006) and a Bayesian Skyline population growth model (Drummond et al., 2005). Four MCMC chain lengths were run for 80 million generations sampling every 8000th generation. Four independent runs were done to improve convergence and effective sample sizes were assessed using Tracer 1.4 (<http://beast.bio.ed.ac.uk/Tracer>).

3. Results

3.1. DENV-3 recombinants

We obtained 636 genomic sequences from GenBank (Table S1, in the supplemental material) and the final alignment included 10,170 coding nucleotide positions. We detected a significant recombination event, with an exceedingly small probability of 5.7E-18 of random occurrence obtained with multiple comparisons correction by all the algorithms implemented in RDP 3.0 (RDP, GENECONV, BootScan, MaxChi, Chimaera, Siscan and 3Seq). For this single event (Fig. 1) DQ401690 (Indonesia 1982) and AY496879 (Philippines 1997) appeared as recombinants, while AY858038 (Indonesia 1988), AY496877 (Bangladesh 2002) and FN429900 (Malaysia 1999) appeared as surrogate parental taxa.

3.2. Phylogenetic network of recombinants

We sought to characterize these events using different criteria and the dataset from Table 1. We used a phylogenetic network approach to provide an implicit picture of evolutionary relationships where the sets of parallel edges (splits) represented incompatible and ambiguous signals in the dataset. In this analysis, we included the complete monophyletic groups to which the surrogate parental taxa belong. These groups comprised the surrogate parental lineage and were retrieved from a maximum likelihood tree inferred with Garli 2.0 (Zwickl, 2006) and the 636 whole-genome sequences (Fig. S1). Our findings provided statistically significant evidence for recombination as indicated by a PHI test (p -value <0.05) and the split decomposition output showed multiple bands of parallel edges associating PH-1997 and ID-1982 with two different genotypes (Fig. 2A). Nevertheless, these split networks did not conclusively support genotype assignment either to genotype II or I, but suggested that there was evidence for both.

We created a chimera to better scrutinize our findings in a computer simulation, since the comparative data did not portray a strong tree-like structure. It was built *in silico* by joining together genomic regions from DENV-1 and DENV-4 serotypes and using it alongside other DENV sequences in a Split Decomposition analysis (Fig. S1, in the supplemental material). Fig. S2A illustrates the resulting phylogenetic network. The interesting result was that the chimera had a region of split decomposition due to conflicting evolutionary histories involving two diverse ancestral monophyletic groups, as in our real data.

Fig. 1. Results from the automated tests with RDP. The recombinants ID-1982 and PH-1997 were aligned to a schematic representation of the DENV genome positioned at the center. For each recombinant, the internal rectangles indicate the surrogate parental taxa (labeled as in Table 1).

Table 1
Sequences used in this study following the automated detection test results conducted with RDP.

Genbank Acc. no.	Local and year of isolation	Label	Description
AY744678	French Polynesia, 1989	PF-1989	Genotype I
FN429904	Malaysia, 2000	MY-2000	Genotype I
EU081221	Singapore, 2005	SG-2005	Genotype I
AY676353	Thailand, 1987	TH-1987	Genotype II
GU131906	Cambodia, 2003	KH-2003	Genotype II
FJ461337	Vietnam, 2008	VN-2008	Genotype II
JQ411814	Sri Lanka, 1989	LK-1989	Genotype III
FJ898459	Trinidad and Tobago, 2002	TT-2002	Genotype III
GU131964	Colombia, 2006	CO-2006	Genotype III
DQ401690	Indonesia, 1982	ID-1982	Recombinant
AY496879	Philippines, 1997	PH-1997	Recombinant
AY496877	Bangladesh, 2002	BD-2002	Minor parent ^a
AY496871	Bangladesh, 2002	BD-2002a	
AY496873	Bangladesh, 2002	BD-2002b	
AY496874	Bangladesh, 2002	BD-2002c	
FN429913	Malaysia, 2004	MY-2004	Major parent ^a
AY858038	Indonesia, 1988	ID-1988	
FN429899	Malaysia, 1997	MY-1997	
FN429900	Malaysia, 1999	MY-1999	

^a Monophyletic groups.

3.3. Recombination breakpoint locations

We predicted accurately the recombination breakpoints using algorithms for estimating both tree structure and hidden Markov model parameters (HMM) (Husmeier and McGuire, 2003). Fig. 3A and B shows the corresponding probabilistic graphical model applied to four- and six-taxa ID-1982 alignments respectively (Fig. 4A and B for PH-1997 datasets). Most predictions with each HMM methodology were in agreement and confirmed the existence of multiple breakpoints.

Notably, HMM approaches predicted more breakpoints than RDP (Fig. 1), some of which matched gene boundaries except for the NS5 gene, which encompassed a mosaic region. Within this gene, the RdRp motifs were conserved, except for motif P that had a polymorphic residue in a codon (position 3293 Arg → Trp) involved in priming loop stability (88). Interestingly, the only three

sequences that had this amino acid replacement corresponded to the two recombinants and one surrogate parental virus (BD-2002) (Fig. S3, in the supplemental material).

3.4. Effect of recombination on the accuracy of phylogeny

We calculated PHI, applied the network approach and inferred Maximum clade credibility trees using Bayesian methods to study the phylogenetic congruence among different regions. We selected the partitions for each recombinant following Topali's predictions and BD-2002 and ID-1988 as parents for simplicity. We also limited the number of regions to six because of the multiple break-points within the NS5 gene in the case of PH-1997. Figs. 3 and 4C networks evidenced that different regions showed a different evolutionary history that resulted in inconclusive genotyping. In contrast, topologies obtained using a dataset free of recombination signal were not reticulated and no taxon was ambiguously genotyped (Fig. S4, in the supplemental material). PHI *p*-values did not indicate conflicting evolutionary histories (*p*-values >0.05) except for the region c of the ID-1982 dataset, where we obtained statistically significant evidence for recombination (*p*-value of 7.6E–4). This could be a consequence of the topology probabilities around nucleotide position 7470 (Fig. 4A) that did not discriminate between topologies I and III.

Figs. 5 and 6 show the MCC trees for genomes and regions. Taxa associations resembled those of the phylogenetic network analyses, where both recombinants had different genotype assignment: some regions grouped in Genotype II and that others grouped in Genotype I. The TMRCA for regions b, d and f (where BD-2002 and ID-1982 grouped together) were oddly higher (116.9, 213.4 and 364.1 years, respectively) compared with that estimated for the genome (66.1 years) and the remaining regions (μ = 57.60, σ = 3.42) (Fig. 5). Recombination events are known to introduce bias when timing events in a viral genealogy (Awadalla, 2003), but it seems that it was the coalescent method that had problems when taxa sampled 20 years apart grouped together. In contrast, PH-1997 and BD-2002, which were sampled at a shorter time interval, did not show overestimated TMRCA values when grouped together (Fig. 6, regions b and d).

We discarded that sequence conservation or differences in the parsimony informative sites (e.g. a few nucleotide polymorphisms in a very conserved region that could theoretically lead to disparate inferences) could influence the accuracy of genotype estimation.

Fig. 2. Split decomposition and Bayesian inference results on the reconstruction of the evolutionary history of key taxa that evidence DENV-3 mosaicism. First, networks representing a 95% confidence set of topologies estimated in a Split Decomposition analysis with 100 bootstrapping replicates using (A) the complete dataset from Table 1 or (B) the representative genotyped sequences only; black numbers correspond to bootstrapping values and scale bar for distances. Second, Bayesian maximum clade credibility (MCC) trees inferred for (C) the complete dataset from Table 1 and for (D) the representative genotyped sequences only; black numbers correspond to posterior probability estimates, gray numbers correspond to the TMRCA and scale bar for years.

However, we calculated the percentage of segregating sites (S) and nucleotide diversity values (π) (Nei and Li, 1979) in the DNASP 5.0 software (Librado and Rozas, 2009), and values were similar for all regions (S values of $\mu = 14.960$, $\sigma = 0.676$ for ID-1982; $\mu = 15.505$, $\sigma = 1.058$ for PH-1997; π values of $\mu = 0.05180$, $\sigma = 0.00287$ for ID-1982; $\mu = 0.05403$, $\sigma = 0.00489$ for PH-1997). Hence, we argue that a discrepancy between “genealogies” and the “genome tree” could have resulted from differential selective constraints for each region. For example, the analyses using a recombinant-free dataset often yielded a higher TMRCA estimate than the analysis using complete genomes (Fig. S5, in the supplemental material). Lastly, the phylogenetic relationships among the genotypes were uncertain with a mean value of 0.5 ($\sigma = 0.13$) for all posterior probabilities for any other sister group.

4. Discussion

4.1. DENV-3 mosaics

Recombination is one of the main forces underpinning genome diversity, bringing together diverse genomic regions with different

evolutionary histories into a taxon (Awadalla, 2003). Moreover, recombination seems to be quite common in dengue and in a wide range of RNA viruses (Worobey et al., 1999). For example, Holmes et al. (1999) reported a breakpoint in the E gene of a DENV-1 taxon using strains for which contiguous C-prM-E gene sequences were available (2325 kpb). At the same time, Worobey et al. (1999) reported seven DENV strains from the four serotypes having profiles suggestive of genetic exchange strains and also advised on the apparent inadequacy of short sequence fragments to provide a complete picture of recombinant dengue virus phylogenetic relationships. By analyzing 636 genomic coding regions we found two genomes, one from Indonesia isolated in 1982 and another from the Philippines isolated in 1997, with complex mosaic recombination patterns. Although mosaicism has been previously reported for DENV (Perez-Ramirez et al., 2009), evidence for it is scant. For example, Tolou et al. (2001) and Chen et al. (2008) reported taxa with evidence for multiple recombination events using complete DENV-1 genomic sequences. In the former study, the genomic UTRs were inherited from one ancestral lineage, while an internal fragment located in the E-NS1 region originated from another. The latter study documented a recombinant taxon with

Fig. 3. The recombinant taxon ID-1982. (A) Predictions from Husmeier and Wright HMM with a four-taxon alignment; the horizontal axis shows the position in the genome, and the vertical axis represents the posterior probability for the trees shown in the left. The genomic partitions and a schematic representation of the DENV genome accompany the results positioned at the bottom. (B) Predictions from Westesson and Holmes HMM with six-taxon alignment; the grey and black backgrounds correspond to posterior probabilities of the trees shown in the left. (C) Networks representing a 95% confidence set of topologies estimated in a Split Decomposition analysis with 100 bootstrapping replicates. Most taxon names were removed for simplicity and only those of the recombinant and surrogate parents were maintained; scale bar for distances. Numbers next to region labels represent nucleotide positions.

Fig. 4. The recombinant PH-1997. (A) Predictions from Husmeier and Wright HMM with a four-taxa alignment; the horizontal axis shows the position in the genome, and the vertical axis represents the posterior probability for the trees shown in the left. The genomic partitions and a schematic representation of the DENV genome accompany the results positioned at the bottom. (B) Predictions from Westesson and Holmes HMM with six-taxa alignment; the grey and black backgrounds correspond to posterior probabilities of the trees shown in the left. (C) Networks representing a 95% confidence set of topologies estimated in a Split Decomposition analysis with 100 bootstrapping replicates. Most taxon names were removed for simplicity and only those of the recombinant and surrogate parents were maintained; scale bar for distances. Numbers next to region labels represent nucleotide positions.

Fig. 5. Maximum Clade Credibility trees for the ID-1982 dataset. Most taxa names were removed for simplicity and only those of the recombinant and parents were maintained. Black numbers correspond to posterior probability estimates and gray ones correspond to the TMRCA; scale bar in years.

three recombination regions, located within the sequences of the prM-E junction, NS1 and NS3 genes. More recently, [Carvalho et al. \(2010\)](#) reported DENV-1 recombinants using complete genomic sequences, with two breakpoint positions located in the prM-E region and one within the NS5 region.

We discarded BD-2002 as a recombinant taxon because it consistently grouped within genotype II in our analyses and in previous studies ([Islam et al., 2006](#)). Interestingly, in our ML tree it occupies a basal position together with Thai isolates from 1973 and 1987 ([Supplementary Fig. S1](#), in the supplemental material).

in nature in a low circulation rate and eventually emerged in Bangladesh. Given the location of isolation of the recombinants studied here, we assume that the recombination events took place in South East Asia, where the importation and co-circulation of

multiple viral lineages from all serotypes has been documented (Holmes et al., 2009; Raghwani et al., 2011; Wilder-Smith et al., 2005).

Recombination breakpoints for non-segmented [+] ssRNA viruses have been located in regions of the genome encoding for both, non-structural (Santti et al., 1999) and structural proteins (Costa-Mattioli et al., 2003). We showed that ID-1982 and PH-1997 breakpoints matched gene boundaries well, except for the NS5 gene. Recombination breakpoints, falling *between* rather than *within* genes, suggest that negative selection operated against deleterious intra-gene mosaics, as shown for HIV (Simon-Loriere et al., 2010). In the case of NS5, the recombination events appeared to abide to functional constraints, since five conserved motifs of the RdRp catalytic site fell within a unique partition (Figs. 3 and 4A and B) (Argos, 1988; Poch et al., 1989; Yap et al., 2007). The maintenance of these motifs together with their inter-motif α -helices and β -sheets (Yap et al., 2007) may be important for the correct geometric positioning and orientation of the RdRp catalytic site.

4.2. Inter genotype DENV-3 recombination

We reported recombination among sequences from different genotypes. This may generate new phenotypes by mixing virus-encoded determinants of disease severity (Uzcategui et al., 2001). Similarly, Craig et al. (2003) and Aaskov et al. (2007) reported DENV recombinants composed of portions of two different parental genotypes in a single host. However, the scarcity of data on natural recombinants suggests that such events do not generate viable strains as in the case of Hepatitis C virus (HCV) (Kalinina et al., 2002). Nevertheless, our results, together with the growing evidence of natural recombinants occurring among non segmented [+] ssRNA viruses, suggest that junction of gene regions – or functional partitions, as the catalytic sites of NS5 – may increase the likelihood of viable strains.

Our phylogenetic analyses based on distinct genomic regions further demonstrates that the existence of natural DENV recombinant strains in public databases complicates accurate genotyping, especially in the case of DENV control and molecular epidemiology. For instance, DENV genotyping by phylogenetic analysis of different genomic regions was validated in the past (Jarman et al., 2008) and in the case of DENV-3, all genes and the 3' UTR resulted in satisfactory marker genes. However, if recombinants are present, no single phylogeny describes the ancestry of nucleotide sequences under study because genomic regions will have different evolutionary histories (Degnan and Rosenberg, 2006). Besides, we have also shown that recombination further complicates the estimation of evolutionary parameters, such as the TMRCA (Awadalla, 2003). Appropriately, Posada and Crandall (2002) showed that the effect of recombination on phylogenetic inference depends on the relatedness of the sequences and on the length of the different regions involved, with recombination events being harder to detect among closely related lineages.

4.3. Mosaic structure implications for viral biology and replication

If recombination between divergent strains represents an important source of genetic variation, it may be crucial to broaden the knowledge about RNA viral replication mechanisms and, also search for other recombinants to unveil possible recurrent profiles. The detection of mosaic DENV forms implies that the DENV RdRp protein may have experienced multiple strand switching events during replication. We found an amino acid substitution at position 3293 – within the priming loop – only in the recombinants and a surrogate parental (BD-2002) and hence we hypothesized that such viruses may have an altered structure that allowed the RdRp to access and switch frequently between two strands. The priming

loop stability depends on several intra-loop interactions, and the tryptophan residue at 3293 is involved in a hydrogen bond at the center of this structure (Yap et al., 2007). The replacement of this tryptophan (nonpolar side-chain) by an arginine (positively charged side-chain) in this long and potentially flexible loop may lead to conformational changes that could facilitate strand switching. Since the priming loop regulate the access and exit of RNA templates into the catalytic site (Yap et al., 2007), we would argue that such changes may have a role in the recombination events.

Kim and Kao (2001) have shown that a recombinant RdRp from *Bovine viral diarrhea virus* (BVDV) can switch templates to an internal region of a different strand and then continue synthesis on the initial strand. Besides, although RNA synthesis during elongation is thought to be relatively rapid and processive (Jin et al., 2011), DENV RdRp have a higher conformational flexibility upon the transition from initiation to elongation (Selisko et al., 2006), which may favor recombination. However, this process seems to be a by-product of the processivity of RNA polymerases and not a means by which RNA viruses can purge deleterious mutations (Simon-Loriere and Holmes, 2011). Therefore, template switching could generate mainly deleterious changes and dysfunctional RNA genomes and this could help explain why just a handful of DENV recombinants have being documented so far. Nonetheless, AbuBakar et al. (2002) reported the emergence of a DENV-4 genotype IIA in Malaysia following a recombination event from different ancestral lineages; an observation that, in terms of fitness benefits, advocate for a RNA recombination process that mediates genomic arrangement leading to novel adaptive change.

The identification of recombination events concerning a human pathogen is an important discovery that requires further investigation (Tolou et al., 2001). Our evidence for DENV-3 recombinants was inferred using sequences from public databases. However, given the lack of general agreement about the significance of bioinformatics observations, and due to the lack of appropriate experimental systems, it will be appropriate to verify our observation under laboratory conditions. Experiments by Taucher et al. (2010) showed that intergenomic recombination events may occur in the structural region of flaviviruses, but its frequency appears to be very low. In addition, Chuang and Chen (2009) using site-directed mutagenesis showed that the 5' stem-loop conformation of the secondary structure is important in modulating RNA recombination of the Japanese encephalitis virus. Nonetheless, McGee et al. (2011) using flavivirus plasmid constructs highlighted that in the unlikely event of a high level acute co-infection of two distinct yellow fever genomes, the generation of viable flavivirus recombinants is extremely unlikely.

It will be also necessary to address if recombination occurs more frequently during acute human infection, in mosquitoes or both. Studies have provided evidence that DENV could persist in *Aedes* spp. during successive generations through transovarial transmission (Joshi et al., 2002; Khin and Than, 1983). Consequently, mosquito vectors could play an important role in the maintenance of DENV in nature, acting as reservoirs where concomitant infection may help spawn recombinants. Interestingly, flavivirus-related RNA sequences persist integrated in the dsDNA genome of *Aedes* spp. Mosquitoes, which implies DNA synthesis from viral RNA mediated by an endogenous reverse transcriptase activity (Crochu et al., 2004). Those findings indicate complex interactions between the mosquito's nucleic acid metabolism and DENV replication that may also contribute to the generation of recombinant viruses.

4.4. Conclusion

Our approach for exploring recombination rendered two recombinants so far. In the future, we intend to relax our screening

criteria to study other plausible recombination events (e.g. the hits detected by four or five automated algorithms), and include the UTRs, when available. Several issues regarding DENV recombination remain unsolved, particularly concerning the diversity of recombination forms and the mechanism itself. Crucially, recombination may turn out to be a major determinant of how the virus might respond to vaccination efforts given the variety of pathogenic strains in nature (Holmes and Burch, 2000).

Our observations support the idea that natural recombinants of DENV carrying multiple recombination events (i.e., DENV mosaics) do exist. We recommend further studies for testing the hypothesis that a point mutation resulted in the genetic mosaicism observed because inferences based solely on bioinformatics may be misleading in the absence of experimental validation. New results will support a model of homologous recombination, where multiple strand switching may play a role in generating genetic diversity, with implications for [+] ssRNA virus biology.

Acknowledgements

This work was supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) grant 2010/19059-7. PMAZ holds a CNPq-PQ scholarship; AFB and CJV hold FAPESP scholarships 2011/14537-0 and 2011/17071-2, respectively.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.meegid.2013.05.012>.

References

- Aaskov, J., Buzacott, K., Field, E., Lowry, K., Berlioz-Arthaud, A., Holmes, E.C., 2007. Multiple recombinant dengue type 1 viruses in an isolate from a dengue patient. *J. Gen. Virol.* 88, 3334–3340.
- AbuBakar, S., Wong, P.-F., Chan, Y.-F., 2002. Emergence of dengue virus type 4 genotype IIA in Malaysia. *J. Gen. Virol.* 83, 2437–2442.
- Argos, P., 1988. A sequence motif in many polymerases. *Nucleic Acids Res.* 16, 9909–9916.
- Arnold, J.J., Cameron, C.E., 1999. Poliovirus RNA-dependent RNA polymerase (3Dpol) is sufficient for template switching in vitro. *J. Biol. Chem.* 274, 2706–2716.
- Awadalla, P., 2003. The evolutionary genomics of pathogen recombination. *Nat. Rev. Genet.* 4, 50–60.
- Bandelt, H., Dress, A., 1992. Split decomposition: a new and useful approach to phylogenetic analysis of distance data. *Mol. Phylogent. Evol.* 2, 242–252.
- Becher, P., Tautz, N., 2011. RNA recombination in pestiviruses: cellular RNA sequences in viral genomes highlight the role of host factors for viral persistence and lethal disease. *RNA Biol.* 8, 216–224.
- Bharaj, P., Chahar, H.S., Pandey, A., Diddi, K., Dar, L., Guleria, R., Kabra, S.K., Broor, S., 2008. Concurrent infections by all four dengue virus serotypes during an outbreak of dengue in 2006 in Delhi. *India Virol. J.* 5, 1.
- Bruen, T.C., Philippe, H., Bryant, D., 2006. A simple and robust statistical test for detecting the presence of recombination. *Genetics* 172, 2665–2681.
- Carney, J., Daly, J.M., Nisalak, A., Solomon, T., 2011. Recombination and positive selection identified in complete genome sequences of Japanese encephalitis virus. *Arch. Virol.* 151, 75–83.
- Carvalho, S.E.S., Martin, D.P., Oliveira, L.M., Ribeiro, B.M., Nagata, T., 2010. Comparative analysis of American dengue virus type 1 full-genome sequences. *Virus Genes* 40, 60–66.
- Chambers, T.J., Hahn, C.S., Galler, R., Rice, C.M., 1990. Flavivirus genome organization, expression, and replication. *Annu. Rev. Microbiol.* 44, 649–688.
- Chen, S.-P., Yu, M., Jiang, T., Deng, Y.-Q., Qin, C.-F., Han, J.-F., Qin, E.-D., 2008. Identification of a recombinant dengue virus type 1 with 3 recombination regions in natural populations in Guangdong province. *China Arch. Virol.* 153, 1175–1179.
- Cheng, C., Nagy, P., 2003. Mechanism of RNA recombination in carmo-and tobusviruses: evidence for template switching by the RNA-dependent RNA polymerase in vitro. *J. Virol.* 22, 12033–12047.
- Chuang, C.-K., Chen, W.-J., 2009. Experimental evidence that RNA recombination occurs in the Japanese encephalitis virus. *Virology* 394, 286–297.
- Cooper, P., 1974. On the nature of poliovirus genetic recombinants. *J. Gen. Virol.* 5, 41–49.
- Costa-Mattioli, M., Ferré, V., Casane, D., Perez-Bercoff, R., Coste-Burel, M., Imbert-Marcille, B.-M., Andre, E.C.M., Bressollette-Bodin, C., Billaudel, S., Cristina, J., 2003. Evidence of recombination in natural populations of hepatitis A virus. *Virology* 311, 51–59.
- Craig, S., Thu, H.H.M., Lowry, K., Wang, X.-F., Holmes, E.C., Aaskov, J., 2003. Diverse dengue type 2 virus populations contain recombinant and both parental viruses in a single mosquito host. *J. Virol.* 77, 4463.
- Crochu, S., Cook, S., Attoui, H., Charrel, R.N., De Chesse, R., Belhouchet, M., Lemasson, J.-J., de Micco, P., de Lamballerie, X., 2004. Sequences of flavivirus-related RNA viruses persist in DNA form integrated in the genome of *Aedes* spp. mosquitoes. *J. Gen. Virol.* 85, 1971–1980.
- Dahourou, G., Guillot, S., Le Gall, O., Crainic, R., 2002. Genetic recombination in wild-type poliovirus. *J. Gen. Virol.* 83, 3103–3110.
- Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. *PLoS Genet.* 2, e68.
- Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22, 1185–1192.
- Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4, e88.
- Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973.
- Duggal, R., Cuconati, A., Gromeier, M., Wimmer, E., 1997. Genetic recombination of poliovirus in a cell-free system. *Proc. Natl. Acad. Sci. USA* 94, 13786–13791.
- Edgar, R.C., 2004a. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113.
- Edgar, R.C., 2004b. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.
- Felsenstein, J., Churchill, G.A., 1995. A hidden Markov model approach evolution to variation among sites in rate of evolution. *Mol. Biol. Evol.* 13, 93–104.
- Friedman, N., Ninio, M., Pe'er, I., Pupko, T., 2002. A structural EM algorithm for phylogenetic inference. *J. Comput. Biol.* 9, 331–353.
- Gmyl, A.P., Belousov, E.V., Maslova, S.V., Khitrina, E.V., Chetverin, A.B., Agol, V.I., 1999. Nonreplicative RNA recombination in Poliovirus. *J. Virol.* 73, 8958–8965.
- Gould, E.A., Gritsun, T.S., 2006. Taxonomy and evolution of the flaviviridae. In: Kalitzky, M., Borowski, P. (Eds.), *Molecular Biology of the Flavivirus*. Horizon Bioscience, Norfolk, UK.
- Gould, E.A., Moss, S.R., Turner, S.L., 2004. Evolution and dispersal of encephalitic flaviviruses. *Arch. Virol. Suppl.* 18, 65–84.
- Holmes, E.C., Burch, S.S., 2000. The causes and consequences of genetic variation in dengue virus. *Trends Microbiol.* 8, 74–77.
- Holmes, E.C., Worobey, M., Rambaut, A., 1999. Phylogenetic evidence for recombination in dengue virus. *Mol. Biol. Evol.* 16, 405–409.
- Holmes, E.C., Tio, P.H., Perera, D., Muhi, J., Cardoso, J., 2009. Importation and co-circulation of multiple serotypes of dengue virus in Sarawak. *Malaysia Virus Res.* 143, 1–5.
- Husmeier, D., McGuire, G., 2003. Detecting recombination in 4-taxon DNA sequence alignments with bayesian hidden Markov models and Markov chain Monte Carlo. *Mol. Biol. Evol.* 20, 315–337.
- Huson, D.H., Klopper, T.H., 2005. Computing recombination networks from binary sequences. *Bioinformatics* 21 (Suppl. 2), 159–165.
- Islam, M.A., Ahmed, M.U., Begum, N., Chowdhury, N.A., Khan, A.H., Parquet, M.D.C., Bipolo, S., Inoue, S., Hasebe, F., Suzuki, Y., Morita, K., 2006. Molecular characterization and clinical evaluation of dengue outbreak in 2002 in Bangladesh. *Jpn. J. Infect. Dis.* 59, 85–91.
- Jarman, R.G.R., Holmes, E.C., Rodpradit, P., Klungthong, C., Gibbons, R.V., Nisalak, A., Rothman, A.L., Libraty, D.H., Ennis, F.A., Mammen, M.P., Endy, T.P., 2008. Microevolution of Dengue viruses circulating among primary school children in kamphaeng phet. *Thailand J. Virol.* 8, 1–23.
- Jin, Z., Deval, J., Johnson, K.A., Swinney, D.C., 2011. Characterization of the elongation complex of dengue virus RNA polymerase: assembly, kinetics of nucleotide incorporation, and fidelity. *J. Biol. Chem.* 286, 2067–2077.
- Joshi, V., Mourya, D.T., Sharma, R.C., 2002. Persistence of dengue-3 virus through transovarial transmission passage in successive generations of *Aedes aegypti* mosquitoes. *Am. J. Trop. Med. Hyg.* 67, 158–161.
- Kalinina, O., Norder, H., Mukomolov, S., Magnus, L.O., 2002. A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg. *J. Virol.* 76, 4034–4043.
- Karabatsos, N., 1985. International catalogue of arthropod-borne viruses. A.S.f.T.M.a. Hygiene, San Antonio, TX.
- Khin, M., Than, K., 1983. Transovarial transmission of dengue 2 virus by *Aedes aegypti* in nature. *Am. J. Trop. Med. Hyg.* 32, 590–594.
- Kim, M.J., Kao, C., 2001. Factors regulating template switch in vitro by viral RNA-dependent RNA polymerases: implications for RNA-RNA recombination. *Proc. Natl. Acad. Sci. USA* 98, 4972–4977.
- Laille, M., Deubel, V., Sainte-Marie, F.F., 1991. Demonstration of concurrent dengue 1 and dengue 3 infection in six patients by the polymerase chain reaction. *J. Med. Virol.* 34, 51–54.
- Lavergne, A., de Thoisy, B., Lacoste, V., Pascalis, H., Pouliquen, J.-F., Mercier, V., Tolou, H., Dussart, P., Morvan, J., Talarmin, A., Kazanji, M., 2006. Mayaro virus: complete nucleotide sequence and phylogenetic relationships with other alphaviruses. *Virus Res.* 117, 283–290.
- Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452.
- Lindenbach, B.D., Rice, C.M., 2003. Molecular biology of flaviviruses. *Adv. Virus Res.* 59, 23–61.
- Loroño-Pino, M.A., Cropp, C.B., Farfán, J.A., Vorndam, A.V., Rodríguez-Angulo, E.M., Rosado-Paredes, E.P., Flores-Flores, L.F., Beaty, B.J., Gubler, D.J., 1999. Common

- occurrence of concurrent infections by multiple dengue virus serotypes. *Am. J. Trop. Med. Hyg.* 61, 725–730.
- Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefeuve, P., 2010. RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* 26, 2462–2463.
- McGee, C.E., Tsetsarkin, K.A., Guy, B., Lang, J., Plante, K., Vanlandingham, D.L., Higgs, S., 2011. Stability of yellow fever virus under recombinatory pressure as compared with chikungunya virus. *PLoS One* 6, e23247.
- Milne, I., Lindner, D., Bayer, M., Husmeier, D., McGuire, G., Marshall, D.F., Wright, F., 2009. TOPALI v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics* 25, 126–127.
- Montenegro, F., Araújo, D.C., Maria, R., Nogueira, R., Maria, J., Araújo, G.D., Letícia, I., Ramalho, C., Lucia, M., Sá, F.D., Elizabeth, M., Melo, L.D., Castelo, I., Coelho, B., 2006. Concurrent infection with dengue virus type-2 and DENV-3 in a patient from Ceará. *Brazil Mem. Inst. Oswaldo Cruz* 101, 925–928.
- Nagy, P., Simon, A., 1997. New insights into the mechanisms of RNA recombination. *Virology* 9, 1–9.
- Nájera, R., Delgado, E., Pérez-Alvarez, L., Thomson, M.M., 2002. Genetic recombination and its role in the development of the HIV-1 pandemic. *AIDS* 16 (Suppl. 4), S3–S16.
- Nei, M., Li, W.H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* 76, 5269–5273.
- Perez-Ramirez, G., Diaz-Badillo, A., Camacho-Nuez, M., Cisneros, A., Munoz, M.D.L., 2009. Multiple recombinants in two dengue virus, serotype-2 isolates from patients from Oaxaca. *Mexico BMC Microbiol.* 9, 260.
- Poch, O., Sauvaget, I., Delarue, M., Tordo, N., 1989. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J.* 8, 3867–3874.
- Posada, D., Crandall, K.A., 2002. The effect of recombination on the accuracy of phylogeny estimation. *J. Mol. Evol.* 54, 396–402.
- Raghwan, J., Rambaut, A., Holmes, E.C., Hang, V.T., Hien, T.T., Farrar, J., Wills, B., Lennon, N.J., Birren, B.W., Henn, M.R., Simmons, C.P., 2011. Endemic dengue associated with the co-circulation of multiple viral lineages and localized density-dependent transmission. *PLoS Pathog.* 7, e1002064.
- Reiter, J., Pérez-Vilaró, G., Scheller, N., Mina, L.B., Díez, J., Meyerhans, A., 2011. Hepatitis C virus RNA recombination in cell culture. *J. Hepatol.* 55, 777–783.
- Rice, C., Lenches, E., Shin, S., Sheets, R., Strauss, J., 1985. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. *Science* 229.
- Santti, J., Hyypä, T., Kinnunen, L., 1999. Evidence of recombination among enteroviruses. *J. Virol.* 73, 8741–8749.
- Selisko, B., Dutartre, H., Guillemot, J.-C., Debarnot, C., Benarroch, D., Khromykh, A., Desprès, P., Egloff, M.-P., Canard, B., 2006. Comparative mechanistic studies of de novo RNA synthesis by flavivirus RNA-dependent RNA polymerases. *Virology* 351, 145–158.
- Sibold, C., Meisel, H., Krüger, D.H., Labuda, M., Lysy, J., Kozuch, O., Pejcoch, M., Vaheri, A., Plyusnin, A., 1999. Recombination in Tula hantavirus evolution: analysis of genetic lineages from Slovakia. *J. Virol.* 73, 667–675.
- Simon, A.E., Nagy, P.D., 1996. RNA recombination in turnip crinkle virus: its role in formation of chimeric RNAs, multimers, and in 3' -end repair. *Semin. Virol.* 7, 373–379.
- Simon-Loriere, E., Holmes, E.C., 2011. Why do RNA viruses recombine? *Nat. Rev. Microbiol.* 9, 617–626.
- Simon-Loriere, E., Martin, D.P., Weeks, K.M., Negroni, M., 2010. RNA structures facilitate recombination-mediated gene swapping in HIV-1. *J. Virol.* 84, 12675–12682.
- Taucher, C., Berger, A., Mandl, C.W., 2010. A trans-complementing recombination trap demonstrates a low propensity of flaviviruses for intermolecular recombination. *J. virol.* 84, 599–611.
- Thai, K.T.D., Henn, M.R., Zody, M.C., Tricou, V., Nguyen, N.M., Charlebois, P., Lennon, N.J., Green, L., de Vries, P.J., Hien, T.T., Farrar, J., van Doorn, H.R., de Jong, M.D., Birren, B.W., Holmes, E.C., Simmons, C.P., 2012. High-resolution analysis of intrahost genetic diversity in dengue virus serotype 1 infection identifies mixed infections. *J. Virol.* 86, 835–843.
- Tolou, H.J., Couissinier-Paris, P., Durand, J.P., Mercier, V., de Pina, J.J., de Micco, P., Billoir, F., Charrel, R.N., de Lamballerie, X., 2001. Evidence for recombination in natural populations of dengue virus type 1 based on the analysis of complete genome sequences. *J. Gen. Virol.* 82, 1283–1290.
- Uzcategui, N.Y., Camacho, D., Comach, G., Cuello de Uzcategui, R., Holmes, E.C., Gould, E.A., 2001. Molecular epidemiology of dengue type 2 virus in Venezuela: evidence for in situ virus evolution and recombination. *J. Gen. Virol.* 82, 2945–2953.
- Wang, W.-K., Chao, D.-Y., Lin, S.-R., King, C.-C., Chang, S.-C., 2003. Concurrent infections by two dengue virus serotypes among dengue patients in Taiwan. *J. Microbiol. Immunol. Infect.* 36, 89–95.
- Weaver, S.C., Vasilakis, N., 2009. Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. *Infect. Genet. Evol.* 9, 523–540.
- Wilder-Smith, A., Yoksan, S., Earnest, A., Subramaniam, R., Paton, N.I., 2005. Serological evidence for the co-circulation of multiple dengue virus serotypes in Singapore. *Epidemiol. Infect.* 133, 667–671.
- Worobey, M., Rambaut, a., Holmes, E.C., 1999. Widespread intra-serotype recombination in natural populations of dengue virus. *Proc. Natl. Acad. Sci. USA* 96, 7352–7357.
- Yap, T.L., Xu, T., Chen, Y.-L., Malet, H., Egloff, M.-P., Canard, B., Vasudevan, S.G., Lescar, J., 2007. Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-angstrom resolution. *J. Virol.* 81, 4753–4765.
- Zwickl, D.J., 2006. Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets Under the Maximum Likelihood Criterion. The University of Texas, Austin.

Dengue Virus Type 3 Adaptive Changes during Epidemics in São Jose de Rio Preto, Brazil, 2006–2007

Christian Julian Villabona-Arenas¹*, Adriano Mondini²*, Irene Bosch³, Diane Schmitt⁴, Carlos E. Calzavara-Silva⁵, Paulo M. de A Zanotto¹, Maurício L. Nogueira⁶*

1 Laboratório de Evolução Molecular e Bioinformática (LEMB), Departamento de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, Brazil, **2** Laboratório de Saúde Pública. Departamento de Ciências Biológicas. Faculdade de Ciências Farmacêuticas - Universidade Estadual Paulista "Júlio de Mesquita Filho" Araraquara/SP, Brazil, **3** Division of Health Science and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **4** Department of Infectious Disease and Global Health, Cummings School of Veterinary Medicine, Tufts University, North Grafton, Massachusetts, United States of America, **5** Laboratório de Imunologia Celular e Molecular (LICM), Centro de Pesquisas Rene Rachou (CPqRR), Fundação Oswaldo Cruz (Fiocruz), Belo Horizonte, MG, Brazil, **6** Laboratório de Pesquisas em Virologia, Faculdade de Medicina de São José do Rio Preto, São José do Rio Preto, SP, Brazil

Abstract

Global dengue virus spread in tropical and sub-tropical regions has become a major international public health concern. It is evident that DENV genetic diversity plays a significant role in the immunopathology of the disease and that the identification of polymorphisms associated with adaptive responses is important for vaccine development. The investigation of naturally occurring genomic variants may play an important role in the comprehension of different adaptive strategies used by these mutants to evade the human immune system. In order to elucidate this role we sequenced the complete polyprotein-coding region of thirty-three DENV-3 isolates to characterize variants circulating under high endemicity in the city of São José de Rio Preto, Brazil, during the onset of the 2006-07 epidemic. By inferring the evolutionary history on a local-scale and estimating rates of synonymous (dS) and nonsynonymous (dN) substitutions, we have documented at least two different introductions of DENV-3 into the city and detected 10 polymorphic codon sites under significant positive selection ($dN/dS > 1$) and 8 under significant purifying selection ($dN/dS < 1$). We found several polymorphic amino acid coding sites in the envelope (15), NS1 (17), NS2A (11), and NS5 (24) genes, which suggests that these genes may be experiencing relatively recent adaptive changes. Furthermore, some polymorphisms correlated with changes in the immunogenicity of several epitopes. Our study highlights the existence of significant and informative DENV variability at the spatio-temporal scale of an urban outbreak.

Citation: Villabona-Arenas CJ, Mondini A, Bosch I, Schmitt D, Calzavara-Silva CE, et al. (2013) Dengue Virus Type 3 Adaptive Changes during Epidemics in São Jose de Rio Preto, Brazil, 2006–2007. PLoS ONE 8(5): e63496. doi:10.1371/journal.pone.0063496

Editor: Amit Kapoor, Columbia University, United States of America

Received: February 1, 2013; **Accepted:** April 3, 2013; **Published:** May 7, 2013

Copyright: © 2013 Villabona-Arenas et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: CJA holds a FAPESP scholarship (2011/17071-2); PMAZ and MLN hold a CNPq-PQ scholarship, AM holds an INCT - Dengue scholarship. This work was funded by FAPESP process number 10/19059-7 (PMAZ), by INCT-Dengue and PRONEX-Dengue (CNPq to MLN) and FAPESP Grants # 2012/11733-6 and 2011/10458-9 to MLN. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: mnogueira@famerp.br

† These authors contributed equally to this work.

Introduction

Dengue viruses (genus *flavivirus*, family *Flaviviridae*) are responsible for dengue fever and severe dengue (previously known as dengue hemorrhagic fever, DHF and dengue shock syndrome, DSS), they are some of the most important arthropod-borne viral diseases worldwide [1–4]. Increasing incidences of transmission in urban and semi-urban areas in the tropical and sub-tropical countries have promoted dengue to a major public health issue with a hefty economic burden [5–10].

Dengue epidemics are largely attributed to three factors: (i) Increased urbanization, (ii) global dissemination of the major mosquito vectors, *Aedes aegypti* and *Aedes albopictus*, and (iii) the interaction and evolution of the four DENV serotypes [11]. Inadequate sustained vector control may partially explain unsuccessful disease prevention in developing countries however variables related to the viruses, the human host and the environment should also be considered [12]. Vaccine development is currently underway as a strategy for controlling this worldwide

health threat [13–15]. Due to the complexity of DENV dynamics, prevailing serotypes and differences in distribution of viruses within different regions of the globe, serotype and lineage replacement events may play an important role in the evolution of DENV viruses and, consequently must be addressed during vaccine development [16].

Gene genealogy-based DENV evolutionary studies helped characterize the genetic diversity and distribution of different serotypes/genotypes in space and time [17–22]. Also, they are important for exploring the role of selection on particular dengue proteins [23–29]. Given that serotype-specific neutralizing antibodies confer limited, if any, protection from subsequent infection by the other three serotypes, the appropriate choice of nucleotide residues in the variable genomic regions is critical for the design of effective tetravalent vaccines [30,31].

Brazil is the South American country with the highest economic impact of dengue and also accounts for the majority of reported cases in the continent as evidenced by a 27 year-long study [8,32].

Table 1. GenBank and SJRP sequences used for phylogenetic inference.

Country	Year	GenBank Accession	Genotype
French Polynesia (PF)	1989	AY744678	I
French Polynesia (PF)	1994	AY744685	I
Indonesia (ID)	2004	AY858048	I
Indonesia (ID)	1982	DQ401690	II
Thailand (TH)	1973	DQ863638	II
Thailand (TH)	2001	FJ810414	II
Taiwan (TW)	1998	DQ675531	II
Vietnam (VN)	2006	EU660411	II
Brazil/Acre	2004	EF629367	III
Brazil/Acre	2004	EF629366	III
Brazil/Central West	2001	FJ913015	III
Brazil/Northern	2003	FJ850079	III
Brazil/Northern	2004	FJ850083	III
Brazil/Northern	2005	FJ850086	III
Brazil/Northern	2006	FJ850089	III
Brazil/Northern	2007	FJ850092	III
Brazil/Northern	2008	FJ850094	III
Brazil/Rio de Janeiro	2002	EF629369	III
Brazil/RP	2006	GQ868548	III
Brazil/RP	2006	GQ868547	III
Brazil/RP	2006	GQ868546	III
Brazil	2001	FJ898446	III
Brazil	2001	FJ177308	III
Brazil	2003	FJ898447	III
Colombia (CO)	2007	GQ868578	III
Mozambique (MZ)	1985	FJ882575	III
Nicaragua (NI)	1994	FJ882576	III
Puerto Rico (PR)	2006	EU529692	III
Sri Lanka (LK)	1983	GQ199889	III
Sri Lanka (LK)	1997	GQ252674	III
Trinidad & Tobago (TT)	2002	GQ868617	III
Venezuela (VE)	2007	GQ868587	III
Brazil/Rondonia (BR)	2002	EF629370	V
Brazil/SJRP	2006	RP06161	III
Brazil/SJRP	2006	RP06391	III
Brazil/SJRP	2006	RP06411	III
Brazil/SJRP	2006	RP06611	III
Brazil/SJRP	2006	RP06641	III
Brazil/SJRP	2006	RP06781	III
Brazil/SJRP	2006	RP06801	III
Brazil/SJRP	2006	RP06921	III
Brazil/SJRP	2006	RP06951	III
Brazil/SJRP	2006	RP061901	III
Brazil/SJRP	2006	V065491	III
Brazil/SJRP	2006	RP06171	III
Brazil/SJRP	2006	RP06231	III
Brazil/SJRP	2006	RP06561	III
Brazil/SJRP	2006	RP06651	III

Table 1. Cont.

Country	Year	GenBank Accession	Genotype
Brazil/SJRP	2006	RP06791	III
Brazil/SJRP	2006	V061231	III
Brazil/SJRP	2006	RP06541	III
Brazil/SJRP	2007	RP07071	III
Brazil/SJRP	2007	RP07111	III
Brazil/SJRP	2007	RP07031	III
Brazil/SJRP	2007	RP07141	III
Brazil/SJRP	2007	RP07181	III
Brazil/SJRP	2007	RP07191	III
Brazil/SJRP	2007	RP07251	III
Brazil/SJRP	2007	RP07371	III
Brazil/SJRP	2007	RP07411	III
Brazil/SJRP	2007	RP07531	III
Brazil/SJRP	2007	RP07601	III
Brazil/SJRP	2007	RP07611	III
Brazil/SJRP	2007	RP07661	III
Brazil/SJRP	2007	RP07751	III
Brazil/SJRP	2007	RP07081	III

doi:10.1371/journal.pone.0063496.t001

The highest dengue incidence was among young adults from 2000 to 2007. However, in 2006 the highest incidence rate of severe dengue increased dramatically among children and remained high during severe epidemics in 2007 and 2008 in the state of Rio de Janeiro [32–34]. These circumstances highlight the urgent need for continued studies on dengue molecular epidemiology in order to help determine appropriate policies and effective public health campaigns.

The municipality of São José de Rio Preto (SJRP), one of the 645 municipalities within the State of São Paulo according to the Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatística*, IBGE), has a tropical savanna climate (Köppen climate classification), with an annual rainfall of 1410 mm and mean annual temperature of 23.6°C, making it a favorable breeding ground for *Aedes* mosquito vectors. Our previous study on dengue in SJRP [35] using a 399 bp-long portion of the NS5 gene, suggested that DENV-3 was introduced into the city in 2005. Since SJRP is one of the municipalities in São Paulo with the highest number of dengue cases reported yearly according to the Epidemiological Surveillance Center of the state (*Centro de Vigilância epidemiológica Alexandre Vranjac*, CVE). We chose to further characterize the isolated DENV-3 strains from SJRP from 2006 to 2007 to gain insights into the viral epidemiology that may help control disease.

Materials and Methods

Area of study

The city of SJRP is located in the northwestern region of the state of São Paulo, Brazil (20°49'11" S and 49°22'46" W), with a total area of 431 km² and an estimated population of 415,769 inhabitants in 2012 (Data obtained from the IBGE). The infestation by *Aedes* mosquitoes was reported for the first time in 1985 and the first autochthonous cases were reported in 1990, with the introduction of DENV-1 in 1996 (Chiaravalloti-Neto,

Figure 1. Maximum clade credibility (MMC) tree showing the phylogenetic relationships among SRJP clades (light gray boxes) and other Brazilian isolates. Posterior probability values are shown for main nodes only. Light grey text represents localities (sources) for other Brazilian isolates.

doi:10.1371/journal.pone.0063496.g001

Superintendence of Endemic Disease Control, SUCEN: personal communication). Unpublished reports by the *Instituto Adolfo Lutz* (Brazilian Public Health Laboratory) point to the introduction of DENV-2 in 1998. Dengue viruses have been endemic in the municipality for almost ten years [36].

Sequencing genomes

A total of 33 plasma samples were collected from dengue patients during the 2006 and 2007 outbreaks as described elsewhere [35]. Viral RNA was extracted directly from plasma using the QIAamp Viral RNA mini kit (Qiagen). cDNA was synthesized in a 20 μ l reverse transcription reaction containing: 1 μ l of Superscript III Reverse Transcriptase (Invitrogen), 1 μ l of random hexamers (50 ng/ μ l stock), 1 μ l of specific reverse primer 5'AGAACCTGTTGATTCAACAGCAC3' (10 μ M stock) and 5 μ l of template RNA. Viral cDNA were diluted to 800 μ l in DEPC-treated water and 20 μ l were used as template for a 96 specific PCR reactions. Each 10 μ l PCR reaction contained: 3 μ l of template, 0.03 μ l of pfuUltra II polymerase 1 (5 U/ μ l) (Stratagene), 100 mM dNTPs (Applied Biosystems) and 4 μ l of a mixture of forward and reverse primers (0.5 μ M stock) (Table S1). Primers were synthesized with M13 sequence tags (forward primers with 5'GTAAACGACGGCCAGT3' and reverse primers with 5'CAGGAAACAGCTATGACC3') so that PCR amplicons could be sequenced using universal M13 forward and reverse primers. Each PCR reaction produced 96 overlapping amplicons, each 500–900 nucleotides in length, which were subsequently sequenced bidirectionally using the Big Dye chemistry on ABI3730xl DNA sequencers (Applied Biosystems). The 33

obtained genomic sequences were submitted to the Broad Institute genome project and deposited in GenBank (Table S2 shows all GenBank accession numbers and name of all samples). (<http://www.broadinstitute.org/annotation/viral/Dengue/>).

Phylogenetic Inference

The data set includes the obtained 33 genomic sequences from SJRP that were aligned together with 33 DENV-3 worldwide genomic sequences deposited in GenBank (Table 1) using the Muscle 3.8.31 program [37,38]. Visual inspection of the alignment and manual editing were done with the Se-Al v2.0 program (<http://tree.bio.ed.ac.uk/software/seal/>). A Bayesian maximum clade credibility tree was inferred from a set of plausible trees sampled at the stationary phase of four independent Markov Chain Monte Carlo (MCMC) runs with 40 million generations each using a general time reversible (GTR) model of nucleotide substitution [39], a gamma distributed rate variation and a proportion of invariable sites (GTR + Γ + I). A relaxed (uncorrelated lognormal) molecular clock [40] and a Bayesian Skyline coalescent tree were used as priors [41]. Analyses were done with Beast v1.6.1 [42]. Convergence of parameters was assessed using Tracer v1.5 program (<http://tree.bio.ed.ac.uk/software/tracer/>) until all parameters estimates showed Effective Sample Size (ESS) values over 100.

Selection detection analysis

The SJRP dataset was used in a selection analysis to determine whether amino acid positions were subject to negative or positive

Table 2. Non-neutral codons according to the assayed methods (SLAC and REL) at the specified significance levels ($p=0.05$ and Bayes factor= 50, respectively).

Region	Codon	Residue	Scaled dN-dS (SLAC)	Bayes Factor (REL)	Selection
Endoplasmic reticulum anchor for the protein C	106	V ₂₄ M ₁₀		265.737	Positive
Glycoprotein E: central and dimerization domains	355	P ₃₄	−26.589		Negative
Glycoprotein E: central and dimerization domains	404	P ₂₆ L ₇ S ₁		422.066	Positive
Glycoprotein E: central and dimerization domains	449	T ₃₂ I ₁ A ₁		310.144	Positive
Glycoprotein E: immunoglobulin-like domain	581	T ₃₃ L ₁		349.007	Positive
Nonstructural NS1	865	D ₃₄	−35.355		Negative
Nonstructural NS1	893	K ₃₃ L ₁		738.801	Positive
Nonstructural NS1	935	V ₃₄	−30.491		Negative
Nonstructural NS2B	1401	D ₃₄	−35.355		Negative
NS3 serine Protease: Domain S7	1634	Y ₃₄	−35.355		Negative
NS3 serine Protease: ATP binding domain	1829	A ₂₇ V ₆ S ₁		80226.4	Positive
Nonstructural NS4A: Cytoplasmic domain	2124	H ₃₄	−35.355		Negative
Nonstructural NS4A: Lumenal domain	2191	E ₂₉ D ₁ A ₁		553.026	Positive
AdoMet-Mtase	2605	P ₃₄	−30.491		Negative
NS5: Cytoplasmic domain	2862	V ₃₄	−30.491		Negative
NS5: Cytoplasmic domain	2864	E ₂₀ G ₁₄		527.608	Positive
NS5: Cytoplasmic domain	2879	K ₃₂ R ₂		548.683	Positive
NS5: Catalytic domain	3129	P ₃₂ L ₁ S ₁		286.641	Positive

doi:10.1371/journal.pone.0063496.t002

selection pressures. Comparisons were made with the dengue reference strain Philippines H87/57, following Barrero and Mischenko [43] general approach. Both genealogy-based, codon-site models *Single Likelihood Ancestor Counting* (SLAC) and the Random-Effects (REL) available in the HyPhy package (www.datamonkey.org) were used to estimate the nonsynonymous (dN) and synonymous (dS) rates of substitution [44–46]. Polymorphisms were analyzed with the program DNASP v5 [47] and subjected to the Tajima D statistic test, as proposed by Tajima [48] to evaluate deviations from the neutral expectation of molecular evolution [49].

Epitope and immunogenicity prediction

The amino acid sequences coding for the capsid, envelope, NS1, NS2a and NS5 proteins were aligned (<http://multalin.toulouse.inra.fr/multalin/multalin.html>) and a consensus sequence of each protein was submitted to B (<http://www.cbs.dtu.dk/services/BepiPred/>) and T (<http://www.cbs.dtu.dk/services/NetCTL/>) epitope prediction algorithms. The Allele Frequency Net Database (<http://www.allelefreqencies.net/>) was used to select the HLA classes most predominant in the Southeast of Brazil and used to set up the NetCTL server. The mean value of the epitope propensity scores for each sequence was then classified and plotted according to its potential immunogenicity.

Results and Discussion

Complete genome analyzes define distinct DENV3 lineages circulating in SJRP

The availability of high throughput, parallel DNA sequencing allows for comprehensive information on population genetics at the complete viral genome level. In this study, we determined the sequence of 3390 codon-sites of the complete polyprotein from 33 DENV-3 strains isolated during the 2006 and 2007 outbreaks.

The inferred phylogenetic relationship among DENV-3 strains was summarized in a Bayesian maximum credibility tree (MCC) shown in Figure 1. The majority of Brazilian samples grouped with genotype III, as did most taxa from Latin America isolated since 1994 [20,50–52]. Likewise, the Brazilian strain EF629370 grouped within genotype I, as previously described by Schreiber *et al.* [53]. DENV-3 isolates, other than genotype III, have been reported in Colombia [54] and in the Brazilian states of Minas Gerais and Rondonia, likely imported from Asia [55,56]. The successive circulation of different genotypes is of public health concern given that distinct combinations could have different epidemic potentials and risk burdens [57].

The samples from SJRP all fell within genotype III. The MCC tree indicates that SJRP sequences group with Brazilian isolates from different regions in highly supported clades (Figure 1). Clade A included most of SJRP isolates and is more closely related to an isolate from Acre in 2004 (EF629367_BR_2004). Clade B includes four SJRP isolates grouping with (i) a sample from Central West Brazil (FJ913015_BR_2001) and (ii) the Brazilian sample

Figure 2. Per-site and per-gene dN/dS plot. The calculated SLAC dN/dS and REC posterior probability for positive selection ($dN > dS$) are plotted for each site in the upper and lower part of the polyprotein respectively.
doi:10.1371/journal.pone.0063496.g002

FJ898446_BR_2001 of unknown exactly provenance. Clade C includes two SJRP isolates from 2007 that cluster with a sequence from Rio de Janeiro (EF629369_BR_2002) sampled in 2002. Moreover, in clades A and B there was no temporal separation, suggesting that the 2007 isolates apparently originated from previously circulating 2006 lineages. These results suggest distinct virus introductions into SJRP with one lineage becoming more prevalent and experiencing *in situ* evolution (clade A). Co-circulation of multiple DENV-1 lineages has been documented in Asian and other Brazilian cities [58,59].

Succession of DENV-3 genotypes in Brazil

Our data also provides information on the movement of DENV-3 (genotype III) in South America, because sample FJ850079_BR_2002 obtained in the northern part of Brazil grouped with Colombian and Venezuelan sequences suggesting a plausible transmission route.

DENV-3 (genotype III) was isolated for the first time in Brazil from an autochthonous case of DF in December 2000 from Nova Iguaçu, a city located in Rio de Janeiro (RJ) State, which is the second largest metropolitan area in the country, in 2001–02, a large outbreak of DENV-3 occurred in the neighboring city of Rio de Janeiro [60,61]. Following its introduction into RJ, the virus spread to the neighboring State of São Paulo (SP), and eventually to the municipality of SJRP. A previous introduction of DENV-3

took place in Brazil in Rondonia State (Amazon region), but sample EF629370_BR_2002 position in the MCC tree indicates it belongs to genotype V with its sister taxa coming from Asia. Nevertheless, most genotype III isolates from the north of Brazil grouped with the Caribbean strains (samples from Puerto Rico and Trinidad and Tobago), which suggests the existence of preferential routes of spread into Brazil, possibly facilitated by anthropogenic factors [51,62].

The dynamics of viral change

Sequence comparisons using the reference strain Philippines H87/57 (accession no. M93130) as an outgroup, revealed a total of 893 nucleotide changes, of which 132 corresponded to non-synonymous substitutions. The nucleotide diversity π (that is the average number of nucleotide differences per site between two sequences) [63] within SJRP strains was 0.0052 ± 0.001 . Nevertheless, this π value was one order of magnitude lower than that for the complete data set used for phylogenetic inference ($n = 0.038 \pm 2 \times 10^{-5}$), which suggests a genetic bottleneck following the introduction into a new human population, with the virus experiencing rapid *in situ* evolution afterwards. This is a likely scenario given that we sampled from the first documented DENV-3 outbreak in a city where the local human population is expected to be largely susceptible to this serotype. Likewise, the average value of Tajimas' D for all genes was -1.823 ($p < 0.05$)

Table 3. Genetic analysis per genomic region.

Region	Codon Position	Sites	Segregating sites	ETA*	Pi**	K***	NS sites	S Sites	TajimaD
Capsid Protein C	13–342	330	26	26	0.009	2.806	8	18	–1.952
Polyprotein Propeptide	355–612	258	18	18	0.005	1.166	3	15	–2.472
Envelope Glycoprotein M	616–840	225	19	19	0.008	1.829	0	19	–2.052
Glycoprotein E: central and dimerization domains	841–1722	882	88	88	0.010	8.513	15	73	–2.257
Glycoprotein E: immunoglobulin-like domain	1726–2013	288	31	31	0.011	3.234	7	24	–2.033
NS1	2323–3387	1065	87	88	0.008	8.011	17	71	–2.344
NS2A	3406–4029	624	66	66	0.010	6.054	11	55	–2.311
NS2B	4039–4419	381	33	34	0.007	2.765	5	29	–2.382
NS3 Serine Protease	4459–4923	465	30	30	0.006	2.749	3	27	–2.211
DEXDc domain	4987–5811	825	59	60	0.008	2.717	8	52	–2.241
NS4A	6286–6720	435	40	40	0.010	4.146	6	34	–2.078
NS4B	6727–7455	729	66	66	0.009	6.496	8	58	–2.210
AdoMet-MTase	7633–8133	501	34	34	0.008	3.972	1	33	–1.864
RNA-directed RNA polymerase	8218–10158	1941	161	162	0.008	15.130	24	138	–2.340
Full coding region	1–10170	10170	888	893	0.008	83.398	132	761	–2.373

*Total number of mutations.

**Nucleotide diversity.

***Average number of nucleotide differences.

doi:10.1371/journal.pone.0063496.t003

indicating an excess of low frequency polymorphisms, which would be the case in viral populations expanding following a selective sweep.

The Tajima's *D* value near -2 also implies the outcome of negative selection, most likely due to functional constraints imposed on the viral proteins. In agreement with this hypothesis, the overall *dN/dS* value for the entire polyprotein coding region was 0.068 (estimated 95% CI from 0.057 to 0.080) that implied purifying (stabilizing) selection. Zanutto *et al.* [64] have argued that at a coarse-grain, the most common pressure acting on DENV is purifying selection with little evidence of recent adaptive evolution when comparing distantly related sequences in space and time, which would be expected due to synonymous change saturation at the serotype level. Moreover, given that arboviruses have adapted to diverse cell types it is generally considered that most mutations produced are likely to be deleterious [24]. Accordingly, the per-site SLAC *dN/dS* analyses revealed that only eight codons were non-neutral and negatively-selected ($p < 0.05$) (Table 2). Nevertheless, the presence of low frequency polymorphisms could account for adaptive changes, such as the 10 sites under significant positive selection (Bayes factor > 50) that we found with the per-site REL *dN/dS* analyses.

Selected sites were located in both structural (S) and nonstructural (NS) proteins (Figure 2). Among the four non-neutral sites within the envelope protein, codon 355 coding for A Proline located in the central and dimerization domains, was under negative selection suggesting its structural importance. Likewise, positively-selected codons 404 and 581 in the envelope are near sites associated with disulfide bonds, suggesting adaptive changes near structurally important residues. Interestingly, Barrero and Mistchenko [43] also identified codons 581 and 893 (within the nonstructural NS1 gene) as positively-selected. King *et al.* [65] detected a statistically significant positively-selected site in the NS1 gene of DENV-3 genotype II. They also detected selective

pressure in the C and NS4B genes of genotype I and in the E and NS3 genes of genotype II. We also found additional positively-selected codon sites; 106 (Endoplasmic reticulum anchor for the protein Capsid), 2191 (Lumenal domain of NS4A) and 3129 (catalytic domain of the RNA-directed RNA polymerase) that fell within domains that are important for membrane interactions/rearrangements and virus replication.

Polymorphism analysis of individual genomic regions indicated that the NS5 and the envelope proteins varied the most with 24 and 22 replacement changes respectively, followed by NS1 (17), NS2A (11), NS4B (8) and Capsid (8) (Table 3). Changes found in the glycoprotein M, the AdoMet-MTase region and the NS3 serine protease were mostly conservative, while replacements in the capsid were less conservative. Twiddy *et al.* [66] found codon sites 169 and 380 to be under positive selection in DENV-3, with the latter being located within the central and dimerization domains of the glycoprotein, which is involved in cell tropism. Likewise, Rodpothong and Auewarakul [25] reported one codon site in the prM gene and twelve residues in the E gene to be under positive selection. In summary, our findings along with others, suggest that genes other than the envelope may experience adaptive changes following introduction into a susceptible population, which implies a possible role for either host-immunity interaction or vector adaptations that require further study.

Short-term evolution of DENV-3 genotype III

Studies documenting selectively driven evolution in DENV are limited and mostly available for serotype 2 [25,27–29,43,66], further studies to better characterize possible specific adaptive strategies of each serotype are needed. Therefore we looked at allele replacement in DENV 3 (genotype III).

The allele replacement process, from a population genetics standpoint, has an explicit genealogic representation, since frequency increase of an allele will be manifested by appearing

transformation(s) in the same character occurs along that branch; ◆ same transformation occurs in another branch (homoplasy) and transformation(s) in the same character occur along that branch.
doi:10.1371/journal.pone.0063496.g003

deeper as a synapomorphy in a viral genealogy [67]. Therefore, a synapomorphy indicates a fixed allele in the clade defined by it. To clarify this, we coded each of the ten positively-selected residues as a set of terminal unordered character states, represented as a single capital letter. The most parsimonious reconstruction sets of changes at each internal state in the phylogeny were calculated with MacClade v4.07 [68] using the MCC tree rooted with the reference strain Philippines H87/57. The reconstructed changes are illustrated in Figure 3. They indicate that positively-selected codons (404, 449, 581, 2191, 2879 and 3129) mapped to internal nodes as synapomorphies defining the genotype III. For example, codon 106 Met to Val, (within Protein C, the most non-conservative region in our study) characterized the SJRP clade A and codon 404 Pro to Leu characterized the SJRP clade B. Codon 893 appeared as an autapomorphy for the reference strain H87/57 (M93130). The synapomorphic changes differentiating clades A and B that we observed could reflect some aspects of the intra-genotypic interactions exposed to the human population [22].

Mondini *et al.* [35] described the pattern of spread of two DENV lineages circulating in SJRP, Brazil, during 2006. Based on 82 NS5 sequences, the phylogenetic analysis indicated that all samples were of DENV-3 and related to strains circulating in Martinique during 2000–2001. We have considerably extended that analysis by further showing that different lineages, closely relate to different Brazilian isolates, were co-circulating while experiencing distinct adaptive changes as indicated by their positively-selected synapomorphies. This finding is of relevance, since in several other viral systems, positively-selected changes are adaptive in nature and indicative of adaptive changes imposed by vectors, animal reservoirs or human hosts (*i.e.*, immune escape and herd immunity) [69–73].

When we observed that DENV-3 sequences, previously clustered in clades A, B and C (Figure 3), we conducted *in silico* analysis to verify the influence of amino acids substitutions in the immunogenicity of capsid, envelope, NS1, NS2a and NS5 proteins. We found that amino acid substitutions diminished the B/T epitope propensity scores mainly in the DENV-3 sequences clustered in clades B and C (Figure 4). These analyses suggest that amino acid substitutions in sequences grouped in clade A increases its immunogenicity. As a consequence, DENV-3 strains grouped in clade A should expose more antigenic epitopes and more potentially recognized/accessed by the human immune system. We expected that the most immunogenic strains would be negatively selected in a given population. In this particular study we found exactly the opposite. Thus, one can hypothesize that the ADE phenomenon, which would be responsible for an increased pathogenicity and replication in host, can also affect the viral dynamics in a population level. We wonder whether this increased immunogenicity of clade A can induce an antibody dependent enhancement (ADE)-like response in the individual leading to high titers of the virus, which can facilitate the viral transmission to the mosquitoes and, consequently, increase this strain basic reproductive rate. Our findings are of relevance given that ADE may be manifested by challenges imposed by increasing viral immunogenic diversity.

While most studies document the broad geographic expansion of dengue, urban outbreaks dynamics deserves investigation because they contribute to a better understanding of the genetic diversity of strains during local transmission; knowledge that could be exploited for antiviral and vaccine development. In our analyses of 33 complete polyprotein-coding regions we revealed distinct lineages and detected polymorphisms sites that correlated with changes in the immunogenicity of several epitopes. The study

Figure 4. Rank of immunogenicity of DENV-3 sequences. Thirty-three DENV-3 sequences were submitted to B and T epitope prediction using bioinformatics servers and further classified according to their immunogenic potential. Sequences previously clustered in B and C clades have epitopes, in envelope, NS1, NS2a and NS5, putatively less immunogenic when compared to the same epitopes of sequences clustered in clade A, except for sequence ACY70777.1. Substitutions in the capsid were irrelevant.
doi:10.1371/journal.pone.0063496.g004

provides fine-grain information about the molecular epidemiology of dengue infection.

Supporting Information

Table S1 Primers used for amplification and sequencing.
(DOC)

References

- Gubler DJ (1998) Dengue and dengue hemorrhagic fever. *Clinical microbiology reviews* 11: 480–496.
- Zanotto PMAGEA (2002) Molecular epidemiology, evolution and dispersal of the genus flavivirus; T L, editor. Norwell: Kluwer Academic Publishers. 167–195 p.
- Guzman M, Halstead S, Artsob H (2010) Dengue: a continuing global threat. *Nature Reviews*.
- Simmons CP, Farrar JJ, Chau N, Van Vinh Chau, Wills B (2012) Dengue. *The New England Journal of Medicine*: 1423–1432.
- Garg P, Nagpal J, Khairnar P, Seneviratne SL (2008) Economic burden of dengue infections in India. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 102: 570–577.
- Halasa Ya, Shepard DS, Zeng W (2012) Economic Cost of Dengue in Puerto Rico. *American Journal of Tropical Medicine and Hygiene* 86: 745–752.
- Luz PM, Grinsztajn B, Galvani aP (2009) Disability adjusted life years lost to dengue in Brazil. *Tropical medicine & international health: TM & IH* 14: 237–246.
- Shepard DS, Coudeville L, Halasa Ya, Zambrano B, Dayan GH (2011) Economic impact of dengue illness in the Americas. *The American journal of tropical medicine and hygiene* 84: 200–207.
- Beauté J, Vong S (2010) Cost and disease burden of dengue in Cambodia. *BMC public health* 10: 521.
- Castañeda-Orjuela C, Díaz H, Alvis-Guzman N, Olate A, Rodriguez H, et al. (2012) Burden of Disease and Economic Impact of Dengue and Severe Dengue in Colombia, 2011. *Value in Health Regional Issues* 1: 123–128.
- Gubler DJ (2011) Dengue, Urbanization and Globalization: The Unholy Trinity of the 21(st) Century. *Tropical medicine and health* 39: 3–11.
- Kyle JL, Harris E (2008) Global spread and persistence of dengue. *Annual review of microbiology* 62: 71–92.
- Guy B, Almond JW (2008) Towards a dengue vaccine: progress to date and remaining challenges. *Comparative immunology, microbiology and infectious diseases* 31: 239–252.
- Harris E, Ennis FA (2012) Defeating dengue: a challenge for a vaccine. *Nature medicine* 18: 1622–1623.
- Halstead SB (2012) Dengue vaccine development: a 75% solution? *Lancet* 380: 1535–1536.
- Sabchareon A, Wallace D, Sirivichayakul C, Limkittikul K, Chanthavanich P, et al. (2012) Protective efficacy of the recombinant, live-attenuated, CYD tetravalent dengue vaccine in Thai schoolchildren: a randomised, controlled phase 2b trial. *Lancet* 380: 1559–1567.
- Costa RL, Voloch CM, Schrago CG (2012) Comparative evolutionary epidemiology of dengue virus serotypes. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases* 12: 309–314.
- Rico-Hesse R (1990) Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology* 174: 479–493.
- Zanotto P, Gould EA, Gao GF, Harvey PH, Holmes EC (1996) Population dynamics of flaviviruses revealed by molecular phylogenies. *Proceedings of the National Academy of Sciences of the United States of America* 93: 548.
- Kochel T, Aguilar P, Felices V, Comach G (2008) Molecular epidemiology of dengue virus type 3 in Northern South America: 2000–2005. *Infection, Genetics and*: 2000–2005.
- Araújo JMG, Nogueira RMR, Schatzmayr HGH, Zanotto PMDa, Bello G (2009) Phylogeography and evolutionary history of dengue virus type 3, *Genetics and Evolution* 9: 716–725.
- Villabona-Arenas CJ, Zanotto PMDA (2011) Evolutionary history of Dengue virus type 4: Insights into genotype phylodynamics. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases*: 1–8.
- Zanotto PM, Gibbs MJ, Gould Ea, Holmes EC (1996) A reevaluation of the higher taxonomy of viruses based on RNA polymerases. *Journal of virology* 70: 6083–6096.
- Holmes EC (2003) Patterns of intra-and interhost nonsynonymous variation reveal strong purifying selection in dengue virus. *Journal of virology* 77: 11296.
- Rodpothong P, Auewarakul P (2012) Positive selection sites in the surface genes of dengue virus: phylogenetic analysis of the interserotypic branches of the four serotypes. *Virus genes* 44: 408–414.
- Schmidt DJ, Pickett BE, Camacho D, Comach G, Xhaja K, et al. (2011) A phylogenetic analysis using full-length viral genomes of South American dengue serotype 3 in consecutive Venezuelan outbreaks reveals novel NS5 mutation. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases*: 1–9.
- Bennett SNS, Holmes EEC, Chirivella M, Rodriguez DM, Beltran M, et al. (2003) Selection-driven evolution of emergent dengue virus. *biology and evolution* 20: 1650–1658.
- Bennett SN, Holmes EC, Chirivella M, Rodriguez DM, Beltran M, et al. (2006) Molecular evolution of dengue 2 virus in Puerto Rico: positive selection in the viral envelope accompanies clade reintroduction. *The Journal of general virology* 87: 885–893.
- Twiddy S (2002) Phylogenetic Relationships and Differential Selection Pressures among Genotypes of Dengue-2 Virus. *Virology* 298: 63–72.
- Khan AM, Miotto O, Nascimento EJM, Srinivasan KN, Heiny aT, et al. (2008) Conservation and variability of dengue virus proteins: implications for vaccine design. *PLoS neglected tropical diseases* 2: e272.
- Danecek P, Lu W, Schein CH (2010) PCP consensus sequences of flaviviruses: correlating variance with vector competence and disease phenotype. *Journal of molecular biology* 396: 550–563.
- San Martín JL, Brathwaite O, Zambrano B, Solórzano JOJO, Bouckennooghe A, et al. (2010) The epidemiology of dengue in the Americas over the last three decades: a worrisome reality. *The American journal of tropical medicine and hygiene* 82: 128–135.
- Saúde SdVe (2007) Informe Epidemiológico da Dengue. Janeiro a Dezembro 2007. Available at: http://portal.saude.gov.br/portal/arquivos/pdf/boletim_dengue_010208.pdf. Accessed 24 Janeiro 2013.
- Saúde SdVe (2008) Informe Epidemiológico da Dengue. Janeiro a Novembro 2008. Available at: http://portal.saude.gov.br/portal/arquivos/pdf/boletim_dengue_janeiro_novembro.pdf. Accessed 24 Janeiro 2013.
- Mondini A, de Moraes Bronzoni RV, Nunes SHP, Chiaravalloti Neto F, Massad E, et al. (2009) Spatio-temporal tracking and phylodynamics of an urban dengue 3 outbreak in Sao Paulo, Brazil. *PLoS neglected tropical* 3: e448.
- Mondini A, Neto FC, Sanches MGY, Lopes JCC (2005) Análise espacial da transmissão de dengue em cidade de porte médio do interior paulista. *Rev Saúde Pública* 39: 444–451.
- Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics* 5: 113.
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* 32: 1792–1797.
- Rodriguez F, Oliver J, Marin A, Medina J (1990) The general stochastic model of nucleotide substitution. *Journal of theoretical biology*.
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS biology* 4: e88.
- Drummond A, Rambaut A, Shapiro B (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and*.
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology* 7: 214.
- Barrera PR, Mistchenko AS (2008) Genetic analysis of dengue virus type 3 isolated in Buenos Aires, Argentina. *Virus research* 135: 83–88.
- Pond SLK, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics (Oxford, England)* 21: 676–679.
- Pond S, Kosakovsky Pond SL, Frost SDW (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular biology and evolution* 22: 1208–1222.
- Pond SLK, Frost SDW (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics (Oxford, England)* 21: 2531–2533.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics (Oxford, England)* 25: 1451–1452.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
- Kimura M (1989) The neutral theory of molecular evolution: Cambridge University Press. 384 p.
- Uzcategui NY (2003) Molecular epidemiology of dengue virus type 3 in Venezuela. *Journal of General Virology* 84: 1569–1575.
- Aquino VH, Anatriello E, Gonçalves PF, DA Silva EV, Vasconcelos PFC, et al. (2006) Molecular epidemiology of dengue type 3 virus in Brazil and Paraguay, 2002–2004. *The American journal of tropical medicine and hygiene* 75: 710–715.

Table S2 GenBank accession numbers and name of all samples.
(XLSX)

Author Contributions

Conceived and designed the experiments: AM PMZ MLN IB. Performed the experiments: DS AM CVA PMZ CECS. Analyzed the data: PMZ IB MLN CVA CECS. Contributed reagents/materials/analysis tools: IB PMZ CECS. Wrote the paper: CVA AM DS MN PMZ CECS.

52. Villabona-Arenas CJ, Miranda-Esquivel DR, Jimenez REO (2009) Phylogeny of dengue virus type 3 circulating in Colombia between 2001 and 2007. *Tropical medicine & international health: TM & IH* 14: 1241–1250.
53. Schreiber M, Holmes E, Ong S (2009) Genomic epidemiology of a dengue virus epidemic in urban Singapore. *Journal of*
54. Usme-Ciro J, Mendez J, Tenorio A, Rey G (2008) Simultaneous circulation of genotypes I and III of dengue virus 3 in Colombia. *Virol J* 10: 1–10.
55. Figueiredo L, Ferreira G (2008) Dengue virus 3 genotype 1 associated with dengue fever and dengue hemorrhagic fever, Brazil. *Emerging Infectious* 14: 314–316.
56. Aquino VH, Amarilla AA, Alfonso HL, Batista WC, Figueiredo LTM (2009) New genotype of dengue type 3 virus circulating in Brazil and Colombia showed a close relationship to old Asian viruses. *PLoS one* 4: e7299.
57. Guzman MG, Alvarez A, Vazquez S, Alvarez M, Rosario D, et al. (2012) Epidemiological studies on dengue virus type 3 in Playa municipality, Havana, Cuba, 2001–2002. *International journal of infectious diseases: IJID: official publication of the International Society for Infectious Diseases* 16: e198–203.
58. Raghwani J, Rambaut A, Holmes EC, Hang VT, Hien TT, et al. (2011) Endemic dengue associated with the co-circulation of multiple viral lineages and localized density-dependent transmission. *PLoS pathogens* 7: e1002064.
59. dos Santos FB, Nogueira FB, Castro MG, Nunes PC, de Filippis AMB, et al. (2011) First report of multiple lineages of dengue viruses type 1 in Rio de Janeiro, Brazil. *Virology journal* 8: 387.
60. Nogueira RM, Miagostovich MP, de Filippis aM, Pereira Ma, Schatzmayr HG (2001) Dengue virus type 3 in Rio de Janeiro, Brazil. *Memórias do Instituto Oswaldo Cruz* 96: 925–926.
61. Casali CGC, Pereira MMRR, Santos L, Passos MNP, Fortes B, et al. (2004) A epidemia de dengue/dengue hemorrágico no município do Rio de Janeiro, 2001/2002. *Rev Soc Bras Med* 37: 296–299.
62. Araújo JMG, Nogueira RMR, Schatzmayr HG, Zanotto PMDa, Bello G (2009) Phylogeography and evolutionary history of dengue virus type 3. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases* 9: 716–725.
63. Nei M (1987) *Molecular evolutionary genetics*. New York: Columbia University Press.
64. Zanotto P, Gould EAE, Gao GFG, Harvey PH, Holmes EC (1996) Population dynamics of flaviviruses revealed by molecular phylogenies. *Proceedings of the* 93: 548.
65. King C-C, Chao D-Y, Chien L-J, Chang G-JJ, Lin T-H, et al. (2008) Comparative analysis of full genomic sequences among different genotypes of dengue virus type 3. *Virology journal* 5: 63.
66. Twiddy SS, Woelk CH, Holmes EC (2002) Phylogenetic evidence for adaptive evolution of dengue viruses in nature. *The Journal of general virology* 83: 1679–1689.
67. Zanotto PM, Kallas EG, de Souza RF, Holmes EC (1999) Genealogical evidence for positive selection in the nef gene of HIV-1. *Genetics* 153: 1077–1089.
68. Maddison DM, WP (2005) *MacClade 4: Analysis of Phylogeny and Character Evolution*. Version 4.08a. Available at: <http://macclade.org>. Accessed 24 Janeiro 2013.
69. Ross H, Rodrigo A (2002) Immune-mediated positive selection drives human immunodeficiency virus type 1 molecular variation and predicts disease duration. *Journal of virology*.
70. Ferguson NM, Donnelly Ca, Anderson RM (1999) Transmission dynamics and epidemiology of dengue: insights from age-stratified sero-prevalence surveys. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 354: 757–768.
71. Sheridan I, Pybus OG, Holmes EC, Klennerman P (2004) High-Resolution Phylogenetic Analysis of Hepatitis C Virus Adaptation and Its Relationship to Disease Progression. *Journal of Virology* 78: 3447–3454.
72. Chen R, Holmes EC (2008) The evolutionary dynamics of human influenza B virus. *Journal of molecular evolution* 66: 655–663.
73. Suzuki Y (2008) Positive selection operates continuously on hemagglutinin during evolution of H3N2 human influenza A virus. *Gene* 427: 111–116.

Viewpoints

Detection of Four Dengue Serotypes Suggests Rise in Hyperendemicity in Urban Centers of Brazil

Christian Julián Villabona-Arenas¹, Jessica Luana de Oliveira¹, Carla de Sousa Capra², Karime Balarini³, Mauricio Loureiro⁴, Celso Ricardo Theoto P. Fonseca³, Saulo Duarte Passos⁴, Paolo Marinho de Andrade Zanotto^{1*}

1 Laboratório de Evolução Molecular e Bioinformática, Departamento de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, São Paulo, Brazil, **2** Laboratório de Saúde Pública, Secretaria da Saúde, Prefeitura Municipal de Guarujá, Guarujá, São Paulo, Brazil, **3** Itapema Laboratório de Análises Clínicas, Guarujá, São Paulo, Brazil, **4** Laboratório de Infecção e Pediatría, Faculdade de Medicina de Jundiaí, Jundiaí, São Paulo, Brazil

Dengue fever is the most common infectious disease transmitted by a mosquito and a major economic and disease burden in endemic countries. The reported number of dengue cases in 2013 evidences the disease's disturbing impact on human health in Brazil: 950,193 clinical cases, 3,749 with severe manifestations, and 201 deaths (data up to July 13, 2013 by the Pan American Health Organization); any of the four serotypes was the etiologic agent in different geographical regions. The epidemiological surveillance center "Prof. Alexandre Vranjac" documented in the state of São Paulo a total of 105,973 cases (data up to May 6, 2013). This state is the major industrial and economic powerhouse of Brazil, and its capital is the largest city in the southern hemisphere. In this state, the municipality of São Paulo has 11,253,503 inhabitants and 7,388 inhabitants/km² and is the country's main international center of business, culture, and tourism. Because overcrowded cities present highly favorable conditions for increased transmission, we are deeply concerned that hyperendemicity (the cocirculation of multiple dengue serotypes) may become established in this largest interconnected urban area in Brazil.

Hyperendemicity was first associated with increased transmission and the emergence of severe dengue in Asia in the 1950s and then in the Caribbean in the 1980s [1]. Nevertheless, a similar epidemiological pattern of cocirculation in the most highly populated urban areas of Brazil has not been reported so far and is of utmost relevance, since distinct genotypes from the four serotypes have been reported in Brazil. Just to recapitulate, in 1986 DENV-1 was introduced in the country, causing outbreaks. The first autochthonous cases of DENV-2 and DENV-3 were detected respectively in 1990 and 2000 [2]. DENV-4 was isolated for the first time in 1982 in a focal epidemic in the northwestern region of the Brazilian Amazon. Later, in 2008, this serotype emerged as an important pathogen during the Brazilian

outbreaks from 2010 to 2011 [3]. Remarkably, in 2011, Bastos et al. [4] detected the simultaneous circulation of all four dengue serotypes in the municipality of Manaus, which has 1,802,014 inhabitants and 158 inhabitants/km² and is located in the middle of the Amazon rain forest, in the northern state of Amazonas.

We expect a change in the epidemiological pattern once endemic hyperendemicity and high infection rates lead to an immune population before adulthood. Hence, children less than 16 years old will be at greater risk for dengue [5]. For example, prior to 2007, the disease affected mostly adults (20- to 40-year-old people). Nevertheless, during the 2007–2008 epidemics, over 53% of the cases affected children under 15 years of age [6]. Consistently, Rodriguez-Barraquer et al. [7] argued, and we agree, that the disease's shift towards younger patients observed in Brazil can be partially explained by the accumulation of immunity against multiple serotypes in older people. Therefore, we now have a situation where dengue infections in children have the potential to become a leading cause of hospitalization and death.

Additionally, we need to take into account that simultaneous or sequential epidemics with different serotypes are a common risk factor associated with severe cases [1]. Severity is possibly due to

antibody-dependent enhancement, even though the risk is reduced after infection with two or more serotypes [8]. Severe dengue illness has been seen mainly in infants in Asia, where the presence of circulating dengue antibodies acquired passively by maternal vertical transmission is a frequently reported risk factor [9]. Nonetheless, in Brazil, the sequential introduction of serotypes has been accompanied by mild cases rather than severe ones. In this respect, Halstead [8] argued that this may reflect a failure of clinicians to perform diagnoses that fulfill the requirements of the WHO case definition or be due to local human genetic resistance.

To corroborate if hyperendemicity is being established in a populated area in the country, we determined whether one serotype or multiple ones caused the 2013 epidemic in some critical localities in the state of São Paulo. In collaboration with our public health authorities, we collected acute-phase sera from suspected dengue patients from the cities of Guarujá (located in the coastal region with 290,752 inhabitants and 2,035 inhabitants/km²) and Jundiaí (located in the mountain range with 370,126 inhabitants and 856 inhabitants/km²) from December 20, 2012 to May 2013 (summer months). Jundiaí in the west and Guarujá in the east (seaside) are adjacent to, and tightly interconnected with, the densely populated municipality

Citation: Villabona-Arenas CJ, de Oliveira JL, Capra CdS, Balarini K, Loureiro M, et al. (2014) Detection of Four Dengue Serotypes Suggests Rise in Hyperendemicity in Urban Centers of Brazil. *PLoS Negl Trop Dis* 8(2): e2620. doi:10.1371/journal.pntd.0002620

Editor: Rebeca Rico-Hesse, Baylor College of Medicine, United States of America

Published: February 27, 2014

Copyright: © 2014 Villabona-Arenas et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) grant 2010/19059-7. CJA and JLD hold the FAPESP scholarships 2011/17071-2 and 2013/10382-8 respectively; PMdAZ holds a CNPq - PQ scholarship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: pzanotto@usp.br

Figure 1. Maximum credibility tree (and its four serotype subtrees) showing the evolutionary relationships among the capsid/premembrane junction sequences of 55 strains. Posterior probability support values are shown over the nodes and TMRCA values below the nodes. Reference sequences abbreviated by serotype (D1–4), genotype (*Am*: American, *AmAs*: American/Asian, *As*: Asian, *Cosm*: Cosmopolitan, *Syl*: Sylvatic, *I*, *II*, *III*, *IV*, or *V*), GenBank accession, ISO code for country, and isolation year. doi:10.1371/journal.pntd.0002620.g001

of São Paulo. We expected that this approach would inform on the cocirculation of viruses in the entire metropolitan area. Viral RNA was extracted from sera of 24 positive samples selected at random (20 from Guarujá and the only four samples that tested positive from Jundiaí), and we amplified and sequenced the capsid/premembrane junction that was proposed by Lanciotti et al. [10] for typing DENV. All sequences determined in this study were deposited in GenBank (427 bp; accession KF286626–KF286649). To help in classifying our sequences, a small time-stamped dataset comprising 35 sequences that were representative of both serotypes

and genotypes was retrieved from GenBank and aligned with our sequences. A phylogenetic tree (Figure 1) was built using a Bayesian approach, and our evolutionary estimates matched those of Twiddy et al. [11], which validated our analysis. Sequences from both cities belonged to different serotypes. Remarkably, samples from Guarujá clustered within the four serotypes, while samples from Jundiaí grouped with either DENV-1 or DENV-4. We will argue that these findings corroborate a change in epidemiological pattern accompanying a rise in Brazilian urban hyperendemicity that constitutes a greater challenge for surveillance and

control. Crucially, the presence of two serotypes in the same outbreak may be considered as an important warning for high levels of transmission, since Jundiaí has no significant historic record of epidemics.

Our DENV-1 samples grouped within genotype V, and our DENV-4 samples grouped within genotype II, as they did for strains from those serotypes reported elsewhere in Brazil. Besides, our tree showed closely related viral lineages in both municipalities, which implied viral movement across the metropolitan area. Moreover, DENV-1 samples from Jundiaí fell within divergent clades, suggesting that

there are two different lineages in the city. We consider these results important, since multiple viral lineages cocirculating in densely populated places acted as major dengue transmission hubs as noted in Raghvani et al. [12]. On the other hand, DENV-3 samples grouped with genotype V. DENV-3 genotype III has been the most prevalent genotype in the country, but genotypes I and V have also been associated with Brazilian outbreaks [13]. In this respect, our findings raise the intriguing question of whether our DENV-3 samples represent *in situ* evolution and dissemination from Brazilian regions, or if they are due to a novel introduction from Asia. Most significantly, our DENV-2 samples from Guarujá clustered with the American genotype (samples from American Samoa from 1972 and Venezuela from 1987), while all other Brazilian DENV-2 isolates reported elsewhere in Brazil grouped with sequences from the American/Asian genotype. Asian DENV-2 strains were introduced into the Americas in 1981 and

supposedly displaced the American genotype by the 1990s [14]. We did a Maximum Likelihood analysis with 932 DENV-2 sequences from Genbank, which corroborated our finding that DENV-2 samples grouped with the sequences from Oceania, the Americas, and India isolated from 1969 to 1995 that are commonly referred as the American genotype (data not shown). This finding led us to speculate that: (i) a viral preferential association with asymptomatic or mild disease went unnoticed for a substantial time period, or (ii) there was an introduction from another region, given that Guarujá experiences intense seasonal tourism while being adjacent to Santos, the largest and most important international harbor in Brazil.

The presence of the four serotypes in a single outbreak in one of the most densely populated areas of Brazil is a disturbing finding that has also been documented in Asian countries, particularly in India, which has the largest dengue burden in the world. Recurring dengue epidemics in

that country resulted in the establishment of hyperendemic areas, typically in large, densely populated cities, where most DENV serotypes circulate in a sustained fashion [15]. Roughly two years after the report on the presence of DENV-4 in Manaus (Amazon), we now find the four serotypes cocirculating in the south of the country in the outskirts of the municipality of São Paulo. Therefore, a continued advocacy of long-term prevention and control is imperative. Our concern is that if we ignore the urban hyperendemicity, children will be at greater risk for severe disease [5].

Ethics Statement

Both the Human Research Ethics Committee from the Biomedical Sciences Institute of University of São Paulo and the Research Ethics Committee from the Faculty of Medicine of Jundiaí approved the study; a written informed consent was obtained from all patients.

References

- Gubler DJ (1998) Dengue and dengue hemorrhagic fever. *Clin Microbiol Rev* 11: 480–496.
- Nogueira RM, de Araújo JM, Schatzmayr HG (2007) Dengue viruses in Brazil, 1986–2006. *Rev Panam Salud Publica* 22: 358–63.
- Nunes MR, Faria NR, Vasconcelos HB, Medeiros DB, Silva de Lima CP, et al. (2012) Phylogeography of dengue virus serotype 4, Brazil, 2010–2011. *Emerg Infect Dis* 18: 1858–1864.
- Bastos Mde S, Figueiredo RM, Ramasawmy R, Itapirema E, Gimaque JB, et al. (2012) Simultaneous circulation of all four dengue serotypes in Manaus, State of Amazonas, Brazil in 2011. *Rev Soc Bras Med Trop* 45: 393–394.
- Innis BL (1995) Dengue and dengue hemorrhagic fever. In: Porterfield JS, editor. *Exotic viral infections*. London: Chapman and Hall. pp. 103–140.
- Teixeira MG, Costa MCN, Coelho G, Barreto LM (2008) Recent shifts in age pattern of Dengue Hemorrhagic fever, Brazil. *Emerg Infect Dis* 14: 1663.
- Rodriguez-Barraquer I, Cordeiro MT, Braga C, Souza WV, Marques ET, et al. (2011) From Re-Emergence to Hyperendemicity: The Natural History of the Dengue Epidemic in Brazil. *PLOS Negl Trop Dis* 5: e935. doi:10.1371/journal.pntd.0000935
- Halstead SB (2006) Dengue in the Americas and Southeast Asia: do they differ? *Rev Panam Salud Publica* 20: 407–15.
- Halstead SB (1988) Pathogenesis of dengue: challenges to molecular biology. *Science* 239: 476–481.
- Lanciotti RS, Calisher CH, Gubler DJ, Chang GJ, Vorndam AV (1992) Rapid detection and typing of dengue viruses from clinical samples by using reverse transcriptase-polymerase chain reaction. *J Clin Microbiol* 30: 545–551.
- Twiddy SS, Holmes EC, Rambaut A (2003) Inferring the rate and time-scale of dengue virus evolution. *Mol Biol Evol* 20: 122–129.
- Raghvani J, Rambaut A, Holmes EC, Hang VT, Hien TT, et al. (2011) Endemic dengue associated with the cocirculation of multiple viral lineages and localized density-dependent transmission. *PLOS Pathog* 7: e1002064. doi:10.1371/journal.ppat.1002064
- Aquino VH, Amarilla AA, Alfonso HL, Batista WC, Figueiredo LT (2009) New genotype of dengue type 3 virus circulating in Brazil and Colombia showed a close relationship to old Asian viruses. *PLOS One* 4: e7299. doi:10.1371/journal.pone.0007299
- Cologna R, Armstrong PM, Rico-Hesse R (2005) Selection for virulent dengue viruses occurs in humans and mosquitoes. *J Virol* 79: 853–859.
- Chakravarti A, Arora R, Luxemburger C (2012) Fifty years of dengue in India. *Trans R Soc Trop Med Hyg* 106: 273–282.