New genotype invasion of dengue virus serotype 1 drove massive outbreak in Guangzhou, China

Qinlong Jing[1†*], Sean Wu[2†], Zhengjian He[3†], Lihong Yuan[3], Mengmeng Ma[1], Zhijun Bai[1], Liyun Jiang[1], John Marshall[2], Jiahai Lu[4], Zhicong Yang[1*]

[1] Department of Infectious Diseases, Guangzhou Center for Disease Control and Prevention, Guangzhou, Guangdong, People's Republic of China

[2] Department of Biostatistics and Epidemiology, School of Public Health, University of California, Berkeley, California, United States of America

[3] School of Life sciences and Biopharmaceutics, Guangdong Pharmaceutical University, Guangzhou, Guangdong, People's Republic of China

[4] Department of Medical Statistics and Epidemiology, School of Public Health, Sun Yat-sen University, Guangzhou, Guangdong, People's Republic of China

* Corresponding authors

Email: jingqinlong@126.com (QLJ), yangzc@gzcdc.org.cn (ZCY)

†These authors contributed equally

Running title: New genotype invasion of dengue virus drove outbreak in China

Abstract

Background: Dengue fever is a mosquito-borne infectious disease that has caused major health problems. Variations in dengue virus (DENV) genes are important features of epidemic outbreaks. However, the associations of DENV genes with epidemic scale have not been extensively examined. Here, we assessed new genotype invasion of DENV-1 isolated from Guangzhou in China to evaluate associations with epidemic outbreaks.

Methodology/Principal Findings: We used DENV-1 strains isolated from sera of dengue cases from 2002 to 2016 in Guangzhou for complete genome sequencing. A neighbor-joining phylogenetic tree was constructed to elucidate the genotype characteristics and determine if new genotype invasion correlated with major outbreaks. In our study, a new genotype invasion event was observed during each significant outbreak period in 2002-2003, 2006-2007 and 2013-2014. Genotype II was the main epidemic genotype in 2003 and before. Invasion of genotype I in 2006 caused an unusual outbreak with 765 cases (relative risk (RR)=16.24, 95% confidence interval (CI) =12.41-21.25). At the middle and late stages of the 2013 outbreak, genotype III was introduced to Guangzhou as a new genotype invasion responsible for 37340 cases with RR 541.73 (95%CI=417.78-702.45), after which genotypes I and III began co-circulating. Base mutations occurred after new genotype invasion, and the gene sequence of NS3 protein had the lowest average similarity ratio (99.82%), followed by the gene sequence of E protein (99.86%), as compared to the 2013 strain.

Conclusions/Significance: Genotype replacement and co-circulation of multiple DENV-1 genotypes were observed. New genotype invasion was highly correlated with local unusual outbreaks. In addition to DENV-1 genotype I in the unprecedented outbreak in 2014, new

43 genotype invasion by DENV-1 genotype III occurred in Guangzhou.

44

45 **Keywords:** dengue fever; dengue virus; epidemic determinants; new genotype invasion;
46 outbreak

47

48 Author Summary

49 New genotype invasion of dengue virus highly correlates with the massive outbreaks. In this
50 study, we examined the association of the genotype of dengue virus serorype 1 (DENV-1)
51 from human cases through complete genome sequencing with outbreak scale during 2002 and
52 2016 in Guangzhou, China. It was observed that genotype replacement and co-circulation of
53 multiple genotypes occurred. Most importantly, it indicated that new genotype invasion was
54 highly related with local unusual outbreaks in major outbreak periods in 2002-2003, 2006-2007
55 and 2013-2014. DENV-1 genotype II was the main epidemic genotype in 2003 and before.
56 Invasion of genotype I in 2006 caused an unusual outbreak with 765 cases reported. In addition
57 to genotype I circulation, new genotype invasion by genotype III was the key determinant for
58 the 2014 massive outbreak reaching the highest number of cases with 37340. Furthermore, base
59 mutations appeared after genotype III invasion, and the gene sequence of NS3 protein had the
60 lowest average similarity ratio, followed by the gene sequence of E protein, as compared to the
61 2013 strain.

62

Introduction

Dengue fever (DF), transmitted by the bite of infected *Aedes* mosquitoes, has become the most rapidly spreading arboviral disease in recent decades, with accelerating expansion in affected geographic regions worldwide. Currently, approximately half of the world's population lives in areas at risk of infection. Three hundred ninety million infections and 96 million symptomatic cases occur annually, among which 500,000 individuals suffer from severe dengue, such as dengue hemorrhagic fever or dengue shock syndrome [1, 2]. The disease has huge health and economic effects, with an estimated burden of 25.5 disability-adjusted life years per 100,000 individuals [3]. However, application of the only currently licensed vaccine, CYD-TDV (Dengvaxia; Sanofi Pasteur, Lyon, France) is limited owing to safety issues associated with increased hospitalization risk for individuals who have never been infected with dengue before [4], and no specific interventions to treat the disease have been established.

Dengue virus (DENV), a single-stranded positive-sense RNA virus with an 11-kb genome, contains an open reading frame encoding three structural proteins (C, prM/M, and E) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). The virus can be further classified into four serotypes according to their distinctive antigenicities, i.e., DENV-1–4, and there are diverse genotypes within each serotype. The sequence of each DENV genotype is not always fixed, and frequent variations, recombinations, and lineage turnover or replacement may occur because of selection pressure [5].

In recent decades, DF has become a major threat to public health in Southern China [6]. Frequent and large-scale outbreaks have posed a huge disease burden; in particular, serious outbreaks dominated by DENV-1 have occurred in the last 20 years. Guangzhou has become

the most heavily affected area in China since the 1990s; the number of reported cases during 2005–2017 accounted for 60.86% of all cases nationally, reaching 81% in 2014 alone [7-9]. Outbreaks with highest incidence rates were all caused by DENV-1, leading to extensive studies of associated factors, such as climate, mosquito density, control measures, and virus serotypes [10-13]. Despite this, numerous scientific issues related to large-scale outbreaks remain unresolved, particularly the cause of the massive outbreak in Guangzhou in 2013 and 2014.

Genetic variations in DENV are important factors contributing to the severity of epidemics. However, the associations of such genetic variations with epidemic scale have not been thoroughly examined [2, 14]. Virus variations and resulting invasion by new genotypes may be key factors driving large-scale outbreaks and the development of severe symptoms and death when other confounding factors, such as changes in mosquito vectors, climate, imported cases, tourism, and trade, are not altered [15]. Previous studies have mostly focused on E gene and prM/M gene fragments to analyze genetic variations in DENV, and few studies have reported analysis of complete genome sequences. Analysis based on the E gene indicated that DENV genes exhibit diverse lineages and geographical distributions. Different serotypes comprise various subgroups called genotypes, which can differ in virus virulence and transmission rate [16]. DENV-1 has five different genotypes (I–V), among which genotypes I, IV, and V are still prevalent, whereas genotypes II and III appear to have become dormant. Phylogenetic analyses have shown that different virus isolates with different lineage features nonetheless clustered within the same genotype [17]. The practicality and development of high-throughput whole-genome sequencing and deep sequencing have enabled application of new analytical methods.

A recent complete genome sequence analysis revealed that DENV-1 could be classified into three genotypes (I, II, and III), in contrast to the results of previous genotyping based on the E gene [18]. Lee et al. also reported that DENV-1 could be classified into three genotypes by complete genome sequencing [19] in an investigation of a historically large-scale outbreak in Singapore during 2013 and 2014. Their findings showed that the outbreak was related to the introduction of a new genotype III; however, the roles of different genotypes in driving outbreaks have not been sufficiently evaluated.

There are three scenarios under which new genotype invasion may occur in a specific area. First, genetic variation worldwide produces new genotypes. Second, genotypes that are prevalent in local areas are transmitted to areas in which there are no such genotypes. Third, genotypes that have been silent for years suddenly emerge. The status of dengue in China is still that of an imported disease, that can trigger local transmissions [20]. Guangzhou has emerged as an important hotspot, with the number of reported cases exceeding half those recorded nationwide. Currently, Guangzhou is still regarded as a nonendemic area for dengue, supported by numerous studies of mathematical and statistical modeling, virus evolution, and epidemiology [10, 19, 20]. In mainland China, particularly Guangzhou, many studies have evaluated serotyping and genotyping based on the E gene. However, sequence analysis based on the complete genome has not been performed deeply, and efforts to elucidate the impact of introduction of new genotypes on outbreaks have been insufficient.

Because DENV-1 is a representative serotype causing dengue outbreaks in Guangzhou, in this study, we analyzed new genotype invasion and the capacity of different genotypes to drive outbreaks by phylogenetic analysis based on DENV-1 complete genome sequences. We also

explored the hypothesis of new genotype invasion as the main driver in major outbreak years. Our findings are expected to provide insights into viral evolutionary dynamics and the potential causes of massive outbreaks, which will help improve prevention and control measures for dengue.

Methods

Data sources and case investigations

The 2001–2005 case data were obtained from the archives of the Guangzhou Center for Disease Control and Prevention (GZCDC), including case questionnaires, epidemiological survey reports, phase analysis reports, and summaries. The 2006–2016 case information was extracted from the National Notifiable Infectious Diseases Reporting Information System of China. Once medical institutes reported suspected cases through the system, the local district CDC staff would conduct face-to-face case investigations to collect data on demographics, disease information, clinical manifestations, and travel history (domestic and international). Serum samples were also collected at this time.

Specimen collection

DENV strains were isolated from serum specimens of reported DF cases and preserved by the GZCDC and Sun Yat-sen University (SYSU). Blood samples (3–5 mL) from the acute phase (within 6 days after the date of onset) were collected from consenting patients by a nurse at the visiting medical institute or field CDC staff and then separated to obtain serum. Sera were stored at −80°C until processing.

151

152 **Serotyping and virus isolation**

153 We followed serotyping and virus isolation assay protocols recommended by the World Health
154 Organization [21]. Briefly, all serum samples were extracted with total viral RNA using a
155 QIAamp viral RNA mini kit (Qiagen, Germany). Next, a TaqMan probe-based real-time
156 polymerase chain reaction (PCR) protocol was employed to determine the serotype. Positive
157 sera were diluted 10-fold with the sample treatment solution at 4°C for 2 h before being
158 inoculated onto an *Aedes albopictus* mosquito (C6/36) cell line for virus isolation. Cultures
159 were passaged no more than three times. The experiments were conducted in laboratories at
160 GZCDC and by our collaborator SYSU in Guangzhou.

161

162 **DENV complete genome sequencing**

163 We selected representative virus strains and variants for complete genome sequencing. Specific
164 primers for amplification and sequencing are listed in Additional File 1 [22], and primer
165 synthesis was performed by a qualified third-party biotech company. The PCR product was
166 purified using a QIAquick gel extraction kit (Qiagen) according to the manufacturer's
167 instructions and then sent to a qualified third-party biotech company for sequencing using an
168 ABI 3730xl DNA analyzer platform (Applied Biosystems, CA, USA). Nucleotide sequences
169 were initially assembled with Lasergene software (version 8.0; DNASTAR Inc., Madison, WI,
170 USA), and continuous sequences were aligned using BioEdit software (version 7.0.5). In
171 addition, selected strains were transferred to a third-party biotech company (BGI, China) for
172 complete genome sequencing by next-generation sequencing.

Phylogenetic analysis

DENV-1 reference strains were downloaded from GenBank, and those with a length greater than 10000 bp and associated metadata containing the year and location the isolate was sampled were included for phylogenetic analysis. Included samples were from different countries and regions worldwide, and the isolation year varied from 1944 to 2016. During complete genome sequence analysis, sequences identified from the same country in the same year with an evolutionary distance of zero were excluded through multiple sequence alignment.

All complete genome sequences were subjected to multiple sequence alignment using Mafft version 7 (https://mafft.cbrc.jp/alignment/software/). We then performed sequence similarity analysis with Mega 7.0 software (https://www.megasoftware.net/) and constructed neighbor-joining phylogenetic trees using Kimura's two-parameter model. The robustness of nodes was assessed with 1000 bootstrap replicates. All reference strains were labeled with name, year of isolation, location of isolation, and GenBank accession number. For strains other than Chinese strains, only one reference strain was randomly selected for phylogenetic analysis when multiple strains were isolated from the same year and the same country and when their evolutionary distance was zero after multiple sequence alignment by the K80 model.

Association of new genotype invasion with epidemic scale in major outbreak years

In this study, outbreak year was defined as the year when the annual number of local cases exceeded 500. From 2001 onwards, major outbreak years in Guangzhou have often appeared in two consecutive years, including 2002 and 2003, 2006 and 2007, and 2013 and 2014.

Therefore, these three periods were considered three distinct outbreak years in the analysis. Here, we investigated the association of new genotype invasion with incidence rate by including new genotype as the study factor and the incidence rate in each outbreak year as the dependent variable, while using the average incidence rate during study years (median incidence rate) as the control. The incidence rates were compared with relative risks (RRs) and 95% confidence intervals (95% CIs) calculated to determine whether the effects of genotype were statistically significant.

Comparison of the capacity of different genotypes to drive DF outbreaks during the same epidemic period

Because 2013 and 2014 were the most important years for DF outbreaks in Guangzhou, we evaluated the ability of genotype III to drive more severe infections in comparable communities during these two periods. Comparable communities that shared genotype III or genotype I outbreaks were selected based on similar sizes of permanent resident populations and local environment types. We then calculated RRs and 95% CIs to determine the capacity of different genotypes to drive the outbreak during the same epidemic period.

Association of different genotypes with outbreaks during 2002–2016

In order to explore differences in the capacities of various genotypes to drive outbreaks in different years, we employed univariate and multivariate linear regression models. Community incidence caused by different genotypes was used as the dependent variable, genotype was used as the independent variable, and population density was used as the adjusting variable.

The models were tested by analysis of variance with regression coefficients confirmed by *t* tests, and factors with *P* values of less than 0.05 were retained in the final model.

Statistical analysis and graphing

We used R (version 3.2.2; the R Foundation for Statistical Computing, Vienna, Austria) to process demographic, epidemiological, and genomic data using dplyr and ape packages. The geographic source and other general information for DENV strains were evaluated using descriptive analyses.

Ethics statement

The research protocol was reviewed and approved by the institutional review boards of both the GZCDC and the School of Public Health, SYSU. Written informed consent was obtained from adult participants (age ≥ 18 years old) or parents or legal guardians of children enrolled in the study (age < 18 years old). Consent was also obtained from children ages 7–18 years old.

Results

In total, 1679 DENV-1 complete genome sequences were included in the phylogenetic analysis, including 97 strains from China and 1582 strains from other countries and regions, such as Southeast Asia, East Asia, South America, Central America, Africa, the Middle East, and Europe (Additional File Table S2).

As shown in Table 1, there were 65 DENV-1 strains identified from Guangzhou since 1991, accounting for 67.01% (65/97) of all Chinese strains included in the analysis, among

239 which 48 (73.85%, 48/65) strains were sequenced from 204 serum specimens by our research
240 team and collaborators.

241

242 **Table 1.** DENV-1 genome sequences in Guangzhou in different years

Year	Number	Proportion (%)
1991	1	1.54
1995	**1**	**1.54**
1999	1	1.54
2002	**3**	**4.62**
2003	2	3.08
2006	**6**	**9.23**
2007	3	4.62
2010	1	1.54
2011	1	1.54
2013	**8**	**12.31**
2014	**30**	**46.15**
2015	2	3.08
2016	6	9.23
Total	65	100.00

243

244

DENV-1 genome genotyping

Phylogenetic analysis using complete genome sequences showed that DENV-1 was generally classified into three genotypes, i.e., genotype I, II, and III (Additional File Figure S1). All three genotypes have been observed in DENV-1 outbreaks in China, and genotype III was believed to be new in China at the time of its detection. Specifically, DENV-1 genotype III was first identified during the large outbreak in 2013–2014 in Guangzhou, demonstrating highest similarity with strains from India (JQ922548/India/2005, JQ917404/India/2009) and Singapore (KM403584/Singapore/2013), and no prior outbreaks of genotype III had been recorded in China.

DENV-1 genome genotyping of Guangzhou isolates

As shown in complete genome sequence analysis of 65 DENV-1 strains from Guangzhou, genotypes I, II, and III have all been found circulating in Guangzhou. Large-scale outbreaks typically occurred in the years when a genotype was introduced for the first time, such as 2002, 2006, 2013, and 2014 (Figures 1 and 2).

Fig 1. Phylogenetic tree of complete genome sequences of Chinese DENV-1 strains. Guangzhou strains are labeled in light blue. Strains of genotypes I, II, and III are highlighted in light cyan, lavender, and light green, respectively.

Fig 2. Numbers of local dengue cases and yearly genotype distributions from 2002 to 2016. a) Number of local cases DF reported in Guangzhou (2002–2016). b) Yearly distribution of

DENV-1 genotypes in Guangzhou (2002–2016). The year is indicated when a new genotype invasion was introduced.

In 2006, DENV-1 genotype I was introduced into Guangzhou as a new genotype invasion, causing a large-scale outbreak in that year; 765 local cases were reported, followed by a recurrent outbreak of 20 cases in 2007. During the following years until 2012, no genotype I outbreaks were observed except for a small-scale outbreak in 2011 (33 cases). However, this genotype re-emerged in 2013, 2014, and 2016.

DENV-1 genotype II was first isolated in Guangzhou in 1991 and triggered the largest epidemic to date in 1995, with 5337 local cases reported. Later outbreaks occurred in 2002 (1422 cases) and 2003 (76 cases). However, there were no large-scale genotype II outbreaks in Guangzhou (76 cases in 2003, 20 cases in 2007, and 59 cases in 2010) after 2002.

DENV-1 genotype III was first introduced into Guangzhou in the form of a new genotype invasion in 2013. In that year, genotypes I and III co-circulated, and 1249 local cases were reported. Subsequently, in 2014, a historically unprecedented epidemic of local cases was observed in Guangzhou, with a total of 37,340 local cases reported. In April 2015, genotype III was isolated again in Guangzhou. Nevertheless, there were no further DENV-1 cases after June 2015 until 2016. Overall, DENV-2 was mainly responsible for the DF outbreak in Guangzhou in 2015.

Association of new genotype invasion with epidemic scale in major outbreak years

The incidence rates triggered by different genotypes of new genotype invasion in different

289 years were all higher than the average rates during the study years (median number of reported
290 local cases from 2001 to 2016). Because our viral isolation procedures were initiated in 2002,
291 we included 2002 as the year of invasion of genotype II. We found that genotype III showed
292 the greatest capacity for driving outbreaks, with an RR value as high as 541.73 (95% CI:
293 417.78–702.45), followed by genotype II, with a RR value of 37.83 (95% CI: 29.02–49.26).
294 The capacity of genotype I was relatively lower. However, during the specific year of invasion
295 (2006–2007), the RR value of genotype I still reached 16.24 (95% CI: 12.41–21.25). In general,
296 the epidemic capacity of new genotype invasion in Guangzhou varied, with genotype III being
297 the strongest, and the risk was 14.32 and 33.36 times higher than those of genotypes I and II,
298 respectively (Table 2).
299
300 **Table 2.** Comparison of new genotype invasion and epidemic scale during different years

Year	**New genotype**	**Number of cases**	**Average permanent resident population (10,000)**	**Incidence rate (/100,000)**	**RR**	**95% CI**
2013–2014	III	38589	1300.37	296.75	541.73	(417.78–702.45)
2006–2007	I	785	882.65	8.89	16.24	(12.41–21.25)
2002–2003	II	1498	722.91	20.72	37.83	(29.02–49.26)
Average level	--	57	1040.55	0.55	1	--

301 Note: Average level is the median number of reported local cases from 2001 to 2016.
302

303 **Comparison of the capacities of different genotypes for driving outbreaks during the**
304 **same epidemic period**

305 Among 48 DENV-1 strains sequenced by our research team, there were 36 (75%) strains
306 isolated from different DF cases with clear background information, including the patient's
307 permanent address in Guangzhou and time of onset of symptoms. These strains included 14
308 strains of genotype I, three strains of genotype II, and 19 strains of genotype III across various
309 years, with one in 2002, one in 2006, three in 2007, two in 2010, one in 2011, five in 2013, 17
310 in 2014, two in 2015, and four in 2016. Because of the extremely large outbreak observed
311 during 2013–2014, this period was selected as the study epidemic period to compare the
312 capacities of different genotypes for driving DF outbreaks.

313 In 2014, there were 14 communities identified with outbreaks of genotype III, including
314 Baiyun, Beijing, Dadong, Guangta, Jinsha, Licheng, Liurong, Meihuacun, Nancun, Shadong,
315 Shiweitang, Tangjing, Tongde, and Wushan. In contrast, there were only two communities with
316 outbreaks of genotype I, i.e., Dasha and Shayuan, which were both included as the control
317 group in subsequent analyses. Because the size of the permanent resident population and the
318 type of local environment were similar to those of control communities, we selected Tangjing
319 and Liurong as the study groups and compared the capacities of genotypes I and III to drive
320 outbreaks. Our results demonstrated that genotype III showed more driving force than genotype
321 I in dengue outbreaks as a new genotype in 2014, with an RR value of 1.61 (95% CI: 1.47–
322 1.76; Table 3).

323

324 **Table 3.** Comparison of new genotype invasion and previous genotypes driving dengue outbreaks in 2013

325 and 2014.

Year	Genotype	Number of cases	Permanent resident population	Incidence rate (100,000)	RR	95% CI
2014	III	1442	131015	1100.64	1.61	(1.47–1.76)
	I	765	111963	683.26	1	--
2013	III	177	57192	309.48	2.49	(1.89–3.28)
	I	70	56329	124.27	1	--

326 Note: RR, relative risk; 95% CI, 95% confidence interval of RR value.

327

328 In 2013, only Shiweitang was found to be a site of a genotype III outbreak, whereas four
329 communities, including Zhuguang, Zhongnan, Kuangquan, and Jianggao, had genotype I
330 outbreaks. Similarly, when searching for a suitable control community based on the size of the
331 permanent resident population and the type of local environment, we selected Zhuguang to
332 compare the capacity of genotype I for driving outbreaks with that of genotype III in
333 Shiweitang. The results showed that the risk of genotype III as a new genotype for driving
334 outbreaks in 2013 was 2.49 times higher than that of genotype I (95% CI: 1.89–3.28; Table 3).

335

336 **Association of different genotypes with outbreaks over the years**

337 We used a linear regression model to analyze the relationships between the epidemic capacities
338 of different genotypes where the community incidence caused by different genotypes was
339 selected as the dependent variable, the factorial genotype was included as the independent
340 variable with genotype I being the reference, and population density was entered as the

adjusting variable. Both univariate and multivariate analyses demonstrated that genotype III showed a positive correlation and the greatest regression coefficient in magnitude with statistical significance (Table 4). Additionally, there was no statistically significant association between genotypes II and I.

Table 4. Results of linear regression analysis of genotype III in driving the outbreaks

Model	***β***	***Se***	P^{β}	***F***	P^{f}
Model 1	393.00	134.5	0.006	5.50	0.009
Model 2	390.20	138.14	0.008	3.558	0.025

Note: Model 1 is a univariate model (independent variable: genotype); model 2 is based on model 1 with the population density included as the adjusting factor; P^{β} represents the *P* value of the *t* test of the regression coefficient; and P^{f} represents the *P* value of the *F* value test of the regression equation.

Sequence analysis of the genotype III coding region and each protein gene

In this study, genotype III first appeared in October 2013. We compared the coding sequences of the 2013 strain (KX225487) with those of the 2014–2016 strains and found that the average similarity ratio was 99.88%, indicating that base mutations occurred after the genotype invasion. Specifically, base mutations occurred in all three structural proteins and seven nonstructural proteins. The gene sequence encoding NS3 protein had the lowest average similarity ratio (99.82%), followed by the gene sequence encoding E protein (99.86%), suggesting that NS3 and E gene sequences experienced faster mutation after the DENV-1 genotype III invaded Guangzhou (Table 5, Additional File Figure S2).

360

361 **Table 5.** Sequence similarity between the 2013 strain (KX225487) and strains isolated in 2014–2016 (%)

Strains	**Year**	**C**	**E**	**M**	**NS1**	**NS2A**	**NS2B**	**NS3**	**NS4A**	**NS4B**	**NS5**	**CDS**
KX225487	2013	100	100	100	100	100	100	100	100	100	100	100
KT187559	2014	100	99.93	99.55	99.90	100	100	99.84	100	100	99.96	99.93
KT187560	2014	100	99.93	100	99.90	99.85	100	99.89	99.73	99.86	99.96	99.90
KT187561	2014	100	99.93	99.55	99.90	100	100	99.73	100	100	99.96	99.90
KT187562	2014	100	99.80	100	99.90	100	100	99.84	100	100	99.96	99.92
KT827375	2014	100	99.86	100	99.90	100	100	99.89	99.73	100	99.93	99.92
KT827377	2014	100	99.93	100	99.90	100	100	99.89	100	100	99.96	99.95
KX225483	2014	100	99.87	100	100	100	100	99.89	100	100	99.96	99.95
KX225484	2014	100	99.93	99.56	99.91	99.85	100	99.89	100	99.87	99.93	99.91
KX459386	2014	100	99.93	100	99.81	99.54	100	99.84	100	99.87	99.85	99.86
KX459387	2014	100	99.87	100	100	100	100	99.89	100	100	99.96	99.95
KX459388	2014	99.67	99.87	100	99.81	99.54	100	99.78	100	99.87	99.85	99.83
KX459389	2014	100	99.80	100	99.91	100	100	99.89	100	100	99.96	99.93
KX459390	2014	100	100	100	100	100	100	100	100	100	100	100
KX459391	2014	100	99.93	100	99.91	100	100	99.84	100	100	99.96	99.93
KX459392	2014	100	99.80	100	99.81	100	99.75	99.84	100	99.73	99.93	99.87
KX621249	2014	100	99.87	100	99.91	100	100	99.84	100	99.87	99.78	99.87
KT827378	2015	100	99.73	100	99.81	100	100	99.84	100	99.86	99.92	99.88

GZ2016-18488	2016	99.33	99.39	99.1	99.62	99.54	99.23	99.03	98.94	98.92	99.10	99.22

The similarity ratio showed a downward trend each year, with the highest similarity ratios for coding sequences in 2013 and 2014 (average: 99.92%), followed by 2015 (99.88%) and 2016 (lowest, 99.22%). Similar trends were observed for E, NS1, NS3, and NS5 proteins when evaluating specific protein sequences. Specifically, the similarity ratios of the E gene region were 99.90% in 2014, 99.73% in 2015, and 99.39% in 2016; those in the NS1 gene region were 99.91% in 2014, 99.81% in 2015, and 99.62% in 2016; those in the NS3 gene region were 99.87% in 2014, 99.84% in 2015, and 99.03% in 2016; and those in the NS5 gene region were 99.94% in 2014, 99.92% in 2015, and 99.91% in 2016.

Gene regions encoding C, M, NS2A, NS2B, and NS4A proteins were conserved in 2014 and 2015, and the similarity ratios of most sequences reached 100%. However, the sequence similarity ratios of the 2016 strain in either coding region or each protein region (C, E, M, NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) decreased to varying degrees, among which the similarity ratio of the NS4B gene region was the lowest (98.92%), followed by those of the NS4A, NS3, M, NS5, NS2B, C, E, NS2A, and NS1 protein sequences at 98.94%, 99.03%, 99.10%, 99.10%, 99.23%, 99.33%, 99.39%, 99.54%, and 99.62%, respectively.

Discussion

In this study, genotype replacement and co-circulation of multiple genotypes were observed in DENV-1 outbreaks in Guangzhou, China. We found that DENV-1 genotype II was responsible for 2002 outbreaks. However, no large-scale genotype II outbreaks were detected from viral

383 isolate samples in the following years. Additionally, in 2006, genotype I was first identified as
384 a new genotype invasion and was then found to co-circulate with genotype II in 2007. Similar
385 findings were observed until 2013, when the invasion of new genotype III occurred, replacing
386 genotype II to co-circulate with genotype I. Using complete genome sequence analysis and
387 comparative analysis of the epidemic capacity of different genotypes, we found that there were
388 new DENV-1 genotype invasion events in all major outbreak years in Guangzhou, and the
389 appearance of new genotypes was highly correlated with the scale of the outbreak.

390 In terms of the unprecedented large-scale dengue outbreak in Guangzhou in 2014, we
391 observed invasion of a new genotype (genotype III) of DENV-1 in addition to genotype I,
392 which had dominated the epidemics in previous years. This high-intensity outbreak was mainly
393 driven by the new genotype III. Phylogenetic analysis also showed that DENV-1 genotype III,
394 isolated in 2013 and 2014, had high similarity with strains from India (JQ922548/India/2005,
395 JQ917404/India/2009) and Singapore (KM403584/Singapore/2013).

396 New genotype invasion is an important feature of unusual outbreaks in major dengue-
397 endemic regions worldwide. Various serotypes, genotypes, and their lineage clades are
398 different in terms of viral virulence and epidemic capacity [2, 17], with outbreaks characterized
399 by new genotype invasions being the most remarkable. Introduction of new serotypes or
400 genotypes can often change the dominant circulating viral strains in a region. For example, the
401 epidemics in Malaysia in 1993–1995 could be traced back to the invasion of DENV-3 genotype
402 II in Thailand in 1962–1987. The Malaysian DENV-3 isolate used to be genotype I prior to the
403 invasion and was soon replaced by a high-intensity outbreak of DENV-1 and DENV-2
404 cocirculation during 1995–1998. Changes in different serotypes repeatedly caused intense

outbreaks after 2000, with DENV-1 responsible for the most serious epidemics [23-25]. In India, DENV-2 genotype V was gradually replaced by genotype IV from 1967 to 1996, which was accompanied by severe epidemics [26]. By 2003–2004, a phylogenetic analysis revealed that outbreaks were highly related to the invasion of the new DENV-3 genotype III, which eventually took over the previous DENV-2 genotype IV and became the dominant serotype and genotype [27].

Dengue-affected geographic areas are constantly expanding owing to the emergence of new genotypes. For example, DENV-3 genotype III, first identified in the Indian subcontinent, spread to Africa in the 1980s and was further disseminated to Latin America in the 1990s. Notably, the virulence of DENV-3 genotype III changed and tended to be enhanced during geographic dissemination, as demonstrated by statistically significant distribution of mild and severe cases in phylogenetic analysis [28]. In Venezuela, E gene sequence analysis of DENV-3 isolates in the 2000–2001 outbreak showed their likely origin to be a genotype III strain that had invaded from Nicaragua and Panama. This genotype continued to spread in Central America and Mexico and eventually replaced genotype V, which had been epidemic in Venezuela from the 1960s to the 1970s [29]. In Central and South America, genotype invasions occurred more frequently. DENV-3 genotype III spread twice from the Caribbean to Brazil and was introduced to Paraguay at least three times [30], causing serious dengue outbreaks in both countries and surrounding areas. Phylogenetically, Ecuadorian DENV strains were also associated with isolates with Latin American origin [31]. The outbreak of DENV-2 in Puerto Rico originated from the invasion of the new Asian genotype IIIb, and since then, the clade has been cocirculating in the country with another lineage from the Western hemisphere [32]. With

regard to corresponding variations in virulence, previous studies based on the E gene suggested that positive selection occurred at several amino acid positions of the E gene, and such point mutations resulted in not only enhanced transmission but also increased viral virulence.

DENV-1 is most important serotype causing serious outbreaks in China, Southeast Asia, and the South Pacific in recent years. DENV-1 consists of five genotypes (I–V) according to previous phylogenetic analyses based on the E gene, and there are clades of varied sequence features within each genotype [17]. The strain of DENV-1 causing the outbreak in the South Pacific during 1988–1989 was only distantly phylogenetically related to the dominant strains in the region and was much closer related to the American strain, suggesting that the outbreak was caused by the invasion of a new genotype rather than a sudden outbreak of a previous epidemic strain [33]. In 2001, outbreaks of three different genotypes of DENV-1 (I–III) occurred almost at the same time in Myanmar and the South Pacific region; these outbreaks could be traced back to multiple introductions from neighboring regions of Asia [34]. Similarly, there were DENV-1 outbreaks in Hawaii and Tahiti during 2001–2002, and phylogenetic analysis showed that Hawaiian isolates actually originated in Tahiti through invasion of genotype IV [35]. In China, the 2004 DENV-1 outbreak in Zhejiang was related to an imported case of a patient who had traveled to Thailand [36]. In general, dengue outbreaks in China tend to be exclusively caused by imported cases.

Singapore experienced their largest outbreak in history from 2013 to 2014. DENV-1 replaced DENV-2 as the main serotype in circulation, resulting in a total of 40,508 cases, including 22,170 cases in 2013 and 18,338 cases in 2014 with incidence rates as high as 410.6/100,000 and 335.0/100,000, respectively. The outbreak was ultimately confirmed to be

449 caused by the invasion of a DENV-1 genotype III variant [37]. Further analysis of the genetic
450 variation of the new genotype during the epidemic course revealed that there were three
451 different variants in genotypes generated during the local epidemic course in Singapore. These
452 variants exhibited different temporal and spatial distribution patterns with regard to driving the
453 outbreak [19]. In the same year, the largest outbreak of DF was also observed in Taiwan, with
454 a total of 15,732 cases reported, including 136 cases of dengue hemorrhagic fever and twenty
455 deaths, primarily caused by the new genotype of DENV-1 [38]. Thus, DENV-1 genotype III
456 was a key factor of large-scale outbreaks in Southeast Asia, and the successive unprecedented
457 large-scale outbreaks in Singapore and Taiwan during 2013–2014 were both closely related to
458 the invasion of DENV-1 genotype III.

459 Variations in genotypes and their clades have been shown to cause severe dengue
460 epidemics and cases [38]. In Myanmar, 15,361 cases of dengue hemorrhagic fever/dengue
461 shock syndrome and 192 deaths were reported in 2001, and 95% of the cases were caused by
462 DENV-1 [39]. Further phylogenetic studies have shown that the two lineages of the DENV-1
463 genotype I were previously unknown to the region and probably caused by new variations
464 generated from stochastic epidemic events [40]. In 2015, DENV-4 genotype I clade C caused
465 severe cases in southern India, and sequence analysis demonstrated that there were mutations
466 in amino acid sites involved in viral replication and epitope presentation [41].

467 In this study, we demonstrated that DENV-1 new genotype invasion typically caused
468 dengue outbreaks in Guangzhou, particularly in 2006 and 2013–2014, whereas DENV-1
469 genotype II was the main epidemic genotype in 2003 and before. However, after invasion of
470 the DENV-1 genotype I in 2006, Guangzhou soon experienced the largest outbreak of genotype

I, with 765 local cases reported [42]. During the middle and late stages of the epidemic in 2013, DENV-1 genotype III was introduced to Guangzhou as a new genotype invasion. As a result, 1249 local cases were reported, of which 78 cases developed into severe disease, representing the largest outbreak since 2002. DENV-1 genotype III continued to cause outbreaks in 2014 and eventually led to a record-breaking outbreak, with a total of 37,340 local cases reported. This number was over 2.4 times the sum of all cases reported from 1978 to 2013, and 14,000 cases of hospitalization, 308 severe cases, and five deaths were observed. Complete genome sequence analysis of DENV-1 showed that there were two genotypes (I and III) cocirculating in 2014. No significant variations in genotype I were observed. Therefore, this large-scale outbreak was highly associated with genotype III, and the capacity of genotype III for driving outbreaks was stronger than the capacities of genotypes I and II. DENV-1 genotype III strains appeared only in April 2015 and re-emerged in 2016. Moreover, studies have shown that secondary infections played a negligible role in severe cases during the 2014 outbreak [43], suggesting that new genotypes could increase the risk of developing into severe cases.

Because asymptomatic patients and patients with mild disease usually do not seek medical treatment and the patients may visit the hospital at later stages of the disease, specimens in this study were mainly from symptomatic patients, and no asymptomatic individuals (and few patients with mild disease) underwent virus isolation. Therefore, the studied strains may not have represented the entire infected population. In addition, owing to financial constraints, the number of isolated strains and self-sequenced complete genome data in our study were still limited.

Our current findings demonstrated that serotype replacement and variations in genotypes

and clades, particularly new genotype invasion, were important features of large-scale outbreaks of DF. However, future prospective epidemiological and phylogenetic studies are required to further clarify the genetic variations of new genotypes with different genotypes and clades co-circulating in affected areas as well as the epidemic capacities and scales of resulted outbreaks. Moreover, additionally epidemiological studies of severe cases are needed to comprehensively evaluate new genotype invasion and its capacity for driving local epidemics and causing severe disease. Such studies could provide valuable scientific support for prevention and control efforts as well as early detection in dengue-affected areas.

Acknowledgements

We would like to thank all 12 district CDCs in Guangzhou and their staff for participating in this study. We would also like to thank Fuchun Zhang from Guangzhou 8th People's Hospital and Xiaohong Zhou and Xiaoguang Chen from Southern Medical University in China for valuable suggestions.

Competing interests

The authors declare that they have no conflicts of interest.

Consent for publication

All authors approved the final version of this manuscript.

Funding

This work was supported by the Natural Science Foundation of Guangdong Province (grant no. S2013010013637), the Project for Key Medicine Discipline Construction of Guangzhou Municipality (grant no. 2017-2019-04), the Medical Scientific Research Foundation of Guangdong Province (grant no. A2017481), Science and Technology Plan Project of Guangzhou (grant no. 201804010121 and 201904010154), and the Collaborative Innovation Project of Bureau of Science and Technology of Guangzhou Municipality (grant no. 201704020226 and 201803040006). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

QLJ, JHL, and ZCY generated the idea and organized the study; QLJ, SW, ZJH, LHY, and MMM drafted and modified the manuscript. QLJ, ZJH, and MMM conducted field surveys and collected data; QLJ, LYJ, and ZJB conducted the laboratory test; QLJ, SW, ZJH, and JM conducted data analysis.

References

1. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution and burden of dengue. Nature. 2013;496(7446):504-507. doi: 10.1038/nature12060

2. Katzelnick LC, Coloma J, Harris E. Dengue: knowledge gaps, unmet needs, and research priorities. Lancet Infect Dis. 2017;17(3):e88-e100. doi: 10.1016/S1473-3099(16)30473-X

3. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990-2015: a systematic analysis for the Global

538 Burden of Disease Study 2015. Lancet. 2016;388(10053):1603-1658. doi: 10.1016/s0140-
539 6736(16)31460-x
540 4. Yoon IK, Thomas SJ. Encouraging results but questions remain for dengue vaccine. Lancet
541 Infect Dis. 2018;18(2):125-126. doi: 10.1016/S1473-3099(17)30634-5
542 5. Duong V, Simmons C, Gavotte L, Viari A, Ong S, Chantha N, et al. Genetic diversity and
543 lineage dynamic of dengue virus serotype 1 (DENV-1) in Cambodia. Infect Genet Evol.
544 2013;15:59-68. doi: 10.1016/j.meegid.2011.06.019
545 6. Lai S, Huang Z, Zhou H, Anders KL, Perkins TA, Yin W, et al. The changing
546 epidemiology of dengue in China, 1990-2014: a descriptive analysis of 25 years of nationwide
547 surveillance data. BMC Med. 2015;13:100. doi: 10.1186/s12916-015-0336-1
548 7. Sun J, Lu L, Wu H, Yang J, Xu L, Sang S, et al. Epidemiological trends of dengue in
549 mainland China, 2005-2015. Int J Infect Dis. 2017;57:86-91. doi: 10.1016/j.ijid.2017.02.007
550 8. Xiao JP, He JF, Deng AP, Lin HL, Song T, Peng ZQ, et al. Characterizing a large outbreak
551 of dengue fever in Guangdong Province, China. Infect Dis Poverty. 2016;5:44. doi:
552 10.1186/s40249-016-0131-z
553 9. Luo L, Jiang LY, Xiao XC, Di B, Jing QL, Wang SY, et al. The dengue preface to endemic
554 in mainland China: the historical largest outbreak by Aedes albopictus in Guangzhou, 2014.
555 Infect Dis Poverty. 2017;6(1):148. doi: 10.1186/s40249-017-0352-9
556 10. Cheng Q, Jing Q, Spear RC, Marshall JM, Yang Z, Gong P. Climate and the Timing of
557 Imported Cases as Determinants of the Dengue Outbreak in Guangzhou, 2014: Evidence from
558 a Mathematical Model. PLoS Negl Trop Dis. 2016;10(2):e0004417. doi:
559 10.1371/journal.pntd.0004417
560 11. Chen R, Han GZ. Dengue in China: Comprehensive Phylogenetic Evaluation Reveals
561 Evidence of Endemicity and Complex Genetic Diversity. Am J Trop Med Hyg.
562 2016;94(1):198-202. doi: 10.4269/ajtmh.15-0546
563 12. Xiang J, Hansen A, Liu Q, Liu X, Tong MX, Sun Y, et al. Association between dengue
564 fever incidence and meteorological factors in Guangzhou, China, 2005-2014. Environ Res.
565 2017;153:17-26. doi: 10.1016/j.envres.2016.11.009
566 13. Xu L, Stige LC, Chan KS, Zhou J, Yang J, Sang S, et al. Climate variation drives dengue
567 dynamics. Proc Natl Acad Sci U S A. 2017;114(1):113-118. doi: 10.1073/pnas.1618558114

568 14. Liu WJ, Aaskov JG. Fitness peaks of dengue virus populations. PLoS One.
569 2018;13(1):e0189554. doi: 10.1371/journal.pone.0189554
570 15. Vu TT, Holmes EC, Duong V, Nguyen TQ, Tran TH, Quail M, et al. Emergence of the
571 Asian 1 genotype of dengue virus serotype 2 in viet nam: in vivo fitness advantage and lineage
572 replacement in South-East Asia. PLoS Negl Trop Dis. 2010;4(7):e757. doi:
573 10.1371/journal.pntd.0000757
574 16. Cologna R, Armstrong PM, Rico-Hesse R. Selection for virulent dengue viruses occurs in
575 humans and mosquitoes. J Virol. 2005;79(2):853-859. doi: 10.1128/JVI.79.2.853-859.2005
576 17. Cunha Mdos P, Guimaraes VN, Souza M, de Paula Cardoso D, de Almeida TN, de Oliveira
577 TS, et al. Phylodynamics of DENV-1 reveals the spatiotemporal co-circulation of two distinct
578 lineages in 2013 and multiple introductions of dengue virus in Goias, Brazil. Infect Genet Evol.
579 2016;43:130-134. doi: 10.1016/j.meegid.2016.05.021
580 18. Schreiber MJ, Holmes EC, Ong SH, Soh HS, Liu W, Tanner L, et al. Genomic
581 epidemiology of a dengue virus epidemic in urban Singapore. J Virol. 2009;83(9):4163-4173.
582 doi: 10.1128/JVI.02445-08
583 19. Hapuarachchi HC, Koo C, Kek R, Xu H, Lai YL, Liu L, et al. Intra-epidemic evolutionary
584 dynamics of a Dengue virus type 1 population reveal mutant spectra that correlate with disease
585 transmission. Sci Rep. 2016;6:22592. doi: 10.1038/srep22592
586 20. Sang S, Chen B, Wu H, Yang Z, Di B, Wang L, et al. Dengue is still an imported disease
587 in China: A case study in Guangzhou. Infect Genet Evol. 2015. doi:
588 10.1016/j.meegid.2015.03.005
589 21. WHO. Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control: New
590 Edition2009.
591 22. Wu W, Bai Z, Zhou H, Tu Z, Fang M, Tang B, et al. Molecular epidemiology of dengue
592 viruses in southern China from 1978 to 2006. Virol J. 2011;8:322. doi: 10.1186/1743-422X-8-
593 322
594 23. Kobayashi N, Thayan R, Sugimoto C, Oda K, Saat Z, Vijayamalar B, et al. Type-3 dengue
595 viruses responsible for the dengue epidemic in Malaysia during 1993-1994. American Journal
596 of Tropical Medicine & Hygiene. 1999;60(6):904-909.
597 24. Teng AK, Singh S. Epidemiology and new initiatives in the prevention and control of

598 dengue in Malaysia. Dengue Bull. 2001;25:7-12.
599 25. Mohd-Zaki AH, Brett J, Ismail E, L'Azou M. Epidemiology of dengue disease in Malaysia
600 (2000-2012): a systematic literature review. PLoS Negl Trop Dis. 2014;8(11):e3159. doi:
601 10.1371/journal.pntd.0003159
602 26. Singh UB, Maitra A, Broor S, Rai A, Pasha ST, Seth P. Partial nucleotide sequencing and
603 molecular evolution of epidemic causing Dengue 2 strains. Journal of Infectious Diseases.
604 1999;180(4):959-965.
605 27. Dash PK, Man MP, Saxena P, Abhyankar A, Singh CP, Tewari KN, et al. Reemergence
606 of dengue virus type-3 (subtype-III) in India: Implications for increased incidence of DHF &
607 DSS. Virology Journal. 2006;3(1):1-10.
608 28. Messer WB, Gubler DJ, Harris E, Sivananthan K, de Silva AM. Emergence and global
609 spread of a dengue serotype 3, subtype III virus. Emerging Infectious Diseases. 2003;9(7):800-
610 809.
611 29. Uzcategui NY, Comach G, Camacho D, Salcedo M, Cabello de Quintana M, Jimenez M,
612 et al. Molecular epidemiology of dengue virus type 3 in Venezuela. J Gen Virol.
613 2003;84(6):1569-1575. doi: 10.1099/vir.0.18807-0
614 30. Aquino VH, Anatriello E, Goncalves PF, EV DAS, Vasconcelos PF, Vieira DS, et al.
615 Molecular epidemiology of dengue type 3 virus in Brazil and Paraguay, 2002-2004. Am J Trop
616 Med Hyg. 2006;75(4):710-715.
617 31. Regato M, Recarey R, Moratorio G, de Mora D, Garcia-Aguirre L, Gonzalez M, et al.
618 Phylogenetic analysis of the NS5 gene of dengue viruses isolated in Ecuador. Virus Res.
619 2008;132(1):197-200. doi: 10.1016/j.virusres.2007.10.012
620 32. Bennett SN, Holmes EC, Chirivella M, Rodriguez DM, Beltran M, Vorndam V, et al.
621 Molecular evolution of dengue 2 virus in Puerto Rico: positive selection in the viral envelope
622 accompanies clade reintroduction. J Gen Virol. 2006;87(4):885-893. doi: 10.1099/vir.0.81309-
623 0
624 33. Chungue E, Cassar O, Drouet MT, Guzman MG, Laille M, Rosen L, et al. Molecular
625 epidemiology of dengue-1 and dengue-4 viruses. J Gen Virol. 1995;76(7):1877-1884. doi:
626 10.1099/0022-1317-76-7-1877
627 34. Atchareeya A, Berlioz-Arthaud A, Chow V, Endy T, Lowry K, Ninh TU, et al. Sustained

628 transmission of dengue virus type 1 in the Pacific due to repeated introductions of different
629 Asian strains. Virology. 2004;329(2):505-512.
630 35. Imrie A, Zhao Z, Bennett SN, Kitsutani P, Laille M, Effler P. Molecular epidemiology of
631 dengue in the Pacific: introduction of two distinct strains of dengue virus type-1 into Hawaii.
632 Ann Trop Med Parasitol. 2006;100(4):327-336. doi: 10.1179/136485906X105589
633 36. Xu GZ, Dong HJ, Shi NF, Liu SA, Zhou AM, Cheng ZH, et al. An outbreak of dengue
634 virus serotype 1 infection in Cixi, Ningbo, People's Republic of China, 2004, associated with
635 a traveler from Thailand and high density of Aedes albopictus. Am J Trop Med Hyg.
636 2007;76(6):1182-1188.
637 37. Hapuarachchi HC, Koo C, Rajarethinam J, Chong CS, Lin C, Yap G, et al. Epidemic
638 resurgence of dengue fever in Singapore in 2013-2014: A virological and entomological
639 perspective. BMC Infect Dis. 2016;16:300. doi: 10.1186/s12879-016-1606-z
640 38. Shrivastava A, Soni M, Shrivastava S, Sharma S, Dash PK, Gopalan N, et al. Lineage shift
641 of dengue virus in Eastern India: an increased implication for DHF/DSS. Epidemiol Infect.
642 2015;143(8):1599-1605. doi: 10.1017/S0950268814002751
643 39. Thu HM, Lowry K, Myint TT, Shwe TN, Han AM, Khin KK, et al. Myanmar dengue
644 outbreak associated with displacement of serotypes 2, 3, and 4 by dengue 1. Emerging
645 Infectious Diseases. 2004;10(4):593-597.
646 40. Thu HM, Lowry K, Jiang L, Hlaing T, Holmes EC, Aaskov J. Lineage extinction and
647 replacement in dengue type 1 virus populations are due to stochastic events rather than to
648 natural selection. Virology. 2005;336(2):163-172.
649 41. Vaddadi K, Gandikota C, Venkataramana M. Complete genome characterization and
650 evolutionary analysis of serotype-4 associated with severe dengue. Epidemiol Infect.
651 2017;145(7):1443-1450. doi: 10.1017/S0950268817000243
652 42. Luo L, Liang HY, Hu YS, Liu WJ, Wang YL, Jing QL, et al. Epidemiological, virological,
653 and entomological characteristics of dengue from 1978 to 2009 in Guangzhou, China. J Vector
654 Ecol. 2012;37(1):230-240. doi: 10.1111/j.1948-7134.2012.00221.x
655 43. Wang Q, Lei Y, Zhao L, Wen Y, Wang X, Zhang F, et al. Discriminating the
656 primary/secondary infection of dengue fever by the dynamic variation characteristic of
657 antibody response. J Trop Med. 2016;16(5):576-578.

658

659

660 ## Supporting information

661 **Additional File 1: Table S1. Specific primers for amplification and sequencing of dengue**
662 **virus serotype 1.** Total 35 pairs of primers were designed for sequencing DENV-1 genome.

663 **Additional File 2: Table S2. Distribution of DENV-1 genome sequences in different**
664 **countries and regions.** Total 1679 DENV-1 complete genome sequences were analyzed,
665 including 97 strains from China and 1582 strains from other countries and regions.

666 **Additional File 3: Fig S1. DENV-1 complete genome sequence phylogenetic tree.** 1631
667 genome sequences from GenBank and 48 genome sequences from research team were aligned
668 using Mafft software (version 7). Phylogenetic tree was constructed with the neighbor-joining
669 method with Kimura 2-parameter corrections of multiple substitutions using Mega software
670 (version 7.0). Virus strains from China are indicated by read lines.

671 **Additional File 4: Fig S2. Phylogenetic trees for each protein and complete coding**
672 **sequence of DENV-1 genotype III in Guangzhou**. One strain in 2013 and 18 strains from
673 2014 to 2016 of DENV-1 genotype III from Guangzhou were were aligned using Mafft
674 software (version 7). Phylogenetic tree was constructed with the UPGMA method using Mega
675 software (version 7.0).

676

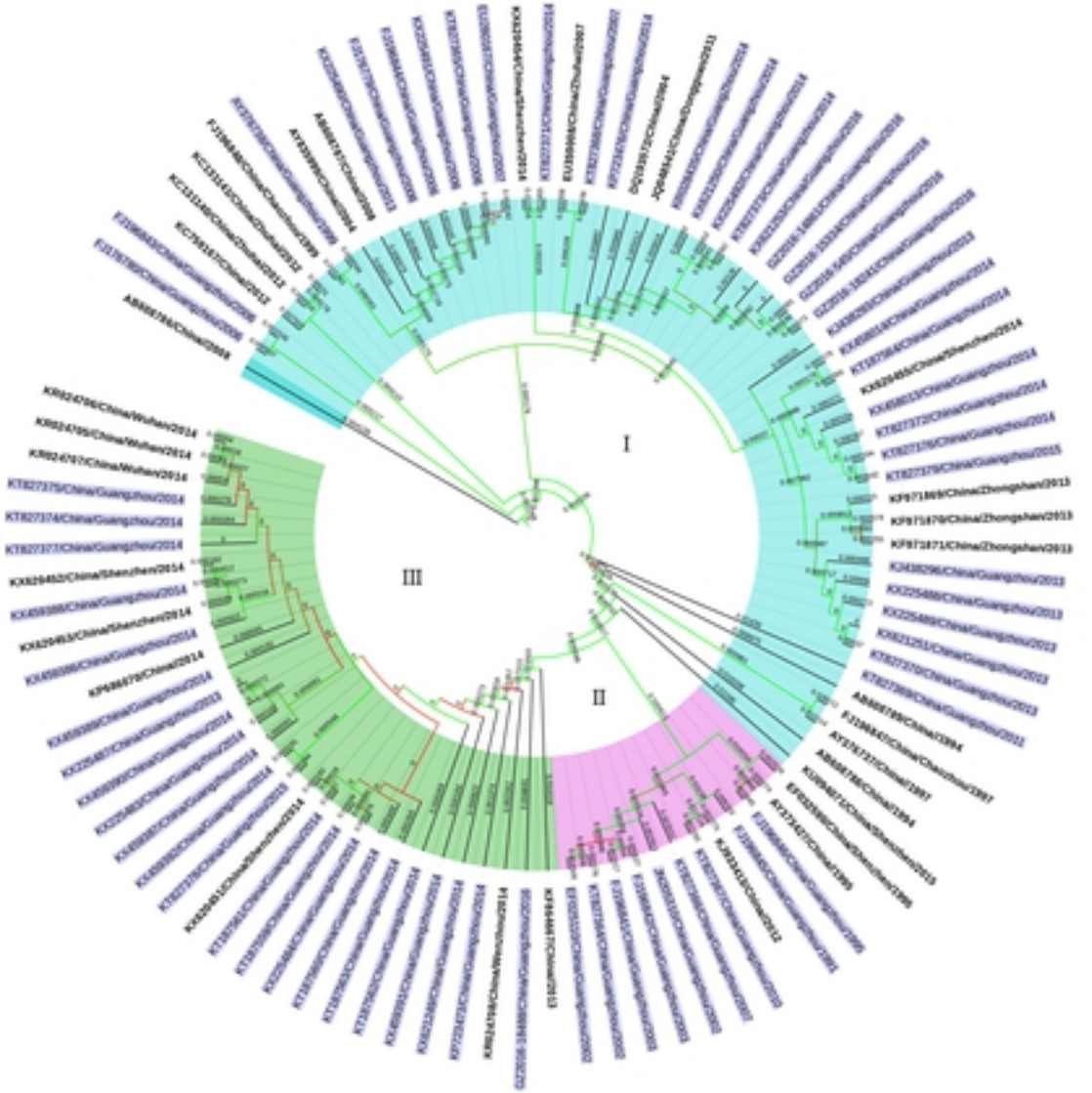

Figure 1

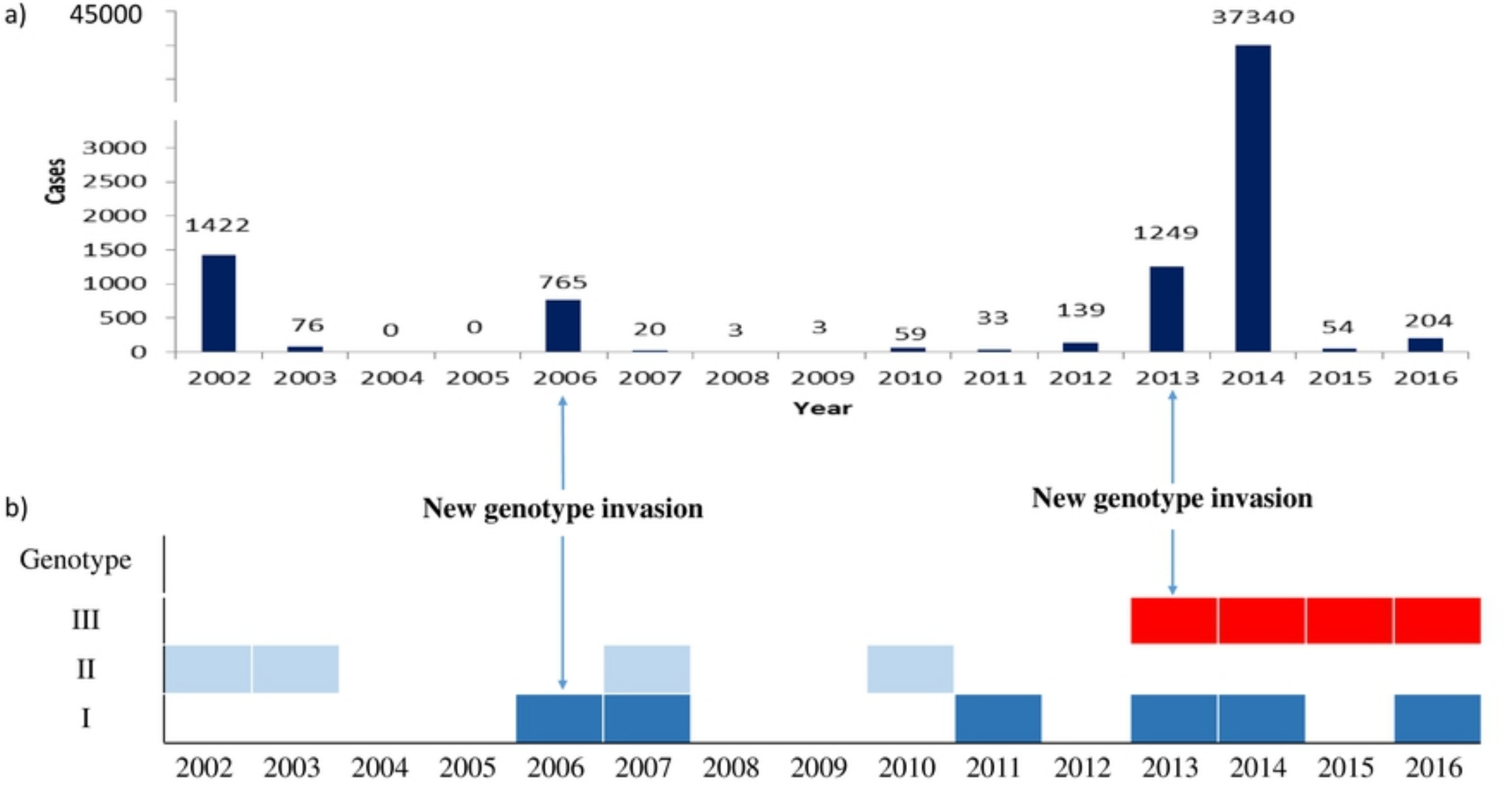
a)
45000
3000
2500
2000
1500
1000
500
0
Cases
1422
76
0
0
765
20
3
3
59
33
139
1249
37340
54
204
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
Year
New genotype invasion
New genotype invasion
b)
Genotype
III
II
I
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016

Figure 2