DISCLAIMER
This paper was submitted to the *Memórias do Instituto Oswaldo Cruz* on 4 April 2017 and was posted to the Fast Track site on 4 April 2017. The information herein is available for unrestricted use, distribution and reproduction provided that the original work is properly cited as indicated by the Creative Commons Attribution licence (CC BY).

RECOMMENDED CITATION
Bonaldo MC, Gómez MM, dos Santos AAC, de Abreu FVS, Ferreira-de-Brito A, de Miranda RM, et al. Genome analysis of yellow fever virus of Brazil ongoing outbreak reveals polymorphisms [Submitted]. Mem Inst Oswaldo Cruz E-pub: 4 Apr 2017. doi: 10.1590/0074-02760170134.

Genome analysis of yellow fever virus of Brazil ongoing outbreak reveals polymorphisms

Myrna C. Bonaldo[1*], Mariela Martínez Gómez[1], Alexandre A. C. dos Santos[1], Filipe Vieira Santos de Abreu[2,3], Anielly Ferreira-de-Brito[2], Rafaella Moraes de Miranda[2], Marcia Gonçalves de Castro[2], Ricardo Lourenço-de-Oliveira[2]

[1]Laboratório de Biologia Molecular de Flavivírus, Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, Brazil

[2]Laboratório de Mosquitos Transmissores de Hematozoários, Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, Brazil

[3]Instituto Federal do Norte de Minas Gerais

Financial support: FAPERJ, CNPq, CAPES, FIOCRUZ

* Corresponding author: mbonaldo@ioc.fiocruz.br

Abstract

The current yellow fever (YF) outbreak in Brazil is the most severe recently reported in the country. It has rapidly spread to areas where YF viral activity have not been observed for more than seventy years and vaccine coverage is almost null. Here, we sequenced the whole YF genome of two naturally infected howler-monkeys (*Alouatta clamitans*) from the Municipality of Domingos Martins, State of Espírito Santo, Brazil. The ongoing-outbreak genome sequences are identical. They clustered in 1E sub-clade (South America I genotype) together with recent Brazilian and Venezuelan strains characterized from infections in humans and non-humans primates. However, we detected eight unique amino acid changes in the viral proteins, which are located in the structural capsid protein (1 change), and the components of viral replicase complex, the NS3 (2 changes) and NS5 (5 changes) proteins, suggesting a potential role in the capacity of viral infection to vertebrate and/or invertebrate hosts and spreading in the ongoing outbreak.

Key words: yellow fever virus, 2017 Brazil outbreak, amino acid changes

Yellow fever virus (YFV) is the prototype member of the genus *Flavivirus*, family Flaviviridae. It is an arbovirus transmitted by the bite of infected mosquitoes in Africa and Americas, causing a disease with a large spectrum of symptoms, from mild disease to severe and deadly hemorrhagic fever in human and New World non-human primates (NHP) (Vasconcelos & Monath 2016). Two main YFV cycles are described: the urban cycle ensured by the domestic mosquito *Aedes* (*Stegomyia*) *aegypti*, currently restricted to Africa, and the wild cycle in which humans are essentially infected during epizooties waves affected NHP, having sylvatic arboreal tree-hole breeding mosquitoes

as vectors (species of *Aedes,* in Africa, and of *Haemagogus* and *Sabethes,* in the Americas). A rural or intermediate cycle may also occurs in zones of emergence recorded in Africa (Monath & Vasconcelos 2015).

YFV is a single-stranded, positive sense RNA virus with a genome of approximately 11 kb. Seven lineages have been identified: five in Africa (West Africa I and II, East Africa, East/Central Africa and Angola), and two in the Americas (South America I and II) (Bryant et al. 2007). Data of molecular and phylogenetic analysis provided evidences that the YFV circulating in the Americas derived from a Western African lineage ancestor emerged in Africa and was imported into the American East coast from West Africa during the slave trade (Bryant et al. 2007; Nunes et al. 2012; Vasconcelos et al. 2004).

The South American I is the most frequent genotype recorded in Brazil (Monath & Vasconcelos 2015; Nunes et al. 2012). Five lineages have been recognized in the South American genotype I, namely IA to IE, which were associated to epidemics recorded during the cyclic expansions and retraction of YFV circulation in Brazil and other tropical American countries (de Souza et al. 2010; Vasconcelos et al. 2004). Since 2008, the lineage ID has been replaced by the emerged lineage IE in Brazil (de Souza et al. 2010; Nunes et al. 2012).

The most severe YFV epidemic reported in Brazil in the recent decades has been reported since late 2016. Until the 10th epidemiological week of 2017, 1,558 cumulative cases with 137 YFV confirmed deaths were reported (COES – Febre Amarela, INFORME – Nº 32/2017). Most importantly, this epidemic has rapidly and alarmingly spread eastward, reaching the most populated Brazilian region where vaccine coverage is minor. Epizooties in NHP and human cases have been diagnosed in states considered YFV-free territories for almost 70 years.

Here, we present the complete genome sequence of two YFV samples collected during the current Brazilian epidemic along with a comparative analysis of recent YFV genome sequences characterized as belonging to the South American genotype I.

Blood samples were obtained from one recently dead and one dying howler-monkeys (*Alouatta clamitans*) found on the Velho Rio farm (20° 17’ 08”S 40° 50’ 15”W), in Areinha, district of Ponto Alto, Municipality of Domingos Martins, State of Espírito Santo, Brazil, in February 20th and 22th 2017. Following centrifugation (2,000 *g* for 10 min), plasma samples were immediately frozen and transported to the laboratory in N_2. Then, plasma samples were screened through RT-PCR. For that, RNA was extracted from 140µL of plasma using the QIAamp Viral RNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s recommendations. RNA was eluted in 60 µl of AVE buffer and stored at -80°C until use. The viral RNA was reverse transcribed applying the High Capacity System (Applied Biosystems) using random hexamers according to the manufacturer’s recommendations. The reverse transcription reaction was carried out at 25°C for 10 min, 37°C for 120 min and 85°C for 5 min. Further, the viral RNA was amplified by conventional PCR using PCR Master Mix (Promega), carried out at 95°C for 2 min, succeeding 30 cycles at 95°C for 1 min, 58°C for 1 min and 72°C for 50 sec; following an extension at 72°C for 5 min. The set of primers utilized in this procedure were: 5'-CTGTGTGCTAATTGAGGTGCATTG-3', and 5'-ATGTCATCAGGCTCTTCTCT- 3’. The YFV infection of the monkeys was confirmed by a specific detection of a single amplicon with the expected YFV amplicon size of 650 bp (Figure 1).

To sequence of the full-length YFV genomes from the positive plasma monkey samples, 12 PCR amplicons were obtained (Supplementary Table 1). At the first step, the viral RNA was reverse transcribed using the Superscript III First-Strand Synthesis

System (Invitrogen) using random hexamers. Alternatively, we generate the first strand cDNA with the reverse primer P11R encoding the 3’UTR end (5´-AGTGGTTTTGTGTTTGTCA-3’) and further processed with YF12F and YF12R to the second strand cDNA synthesis. The cDNA was amplified by conventional PCR using GoTaq Green Master Mix (Promega) according to the manufacturer`s conditions. The thermocycling program set up in a Veriti 96 Well thermocycler (Applied Biosystem) was for regions from (1) to (11): 1 cycle at 95°C for 5 min; 30 cycles at 95°C for 40 sec, at 50°C for 40 sec, and at 72°C for 2 min and finally, 1 cycle at 72°C for 10 min followed by incubation at 4°C. For region (12), we applied 1 cycle at 95°C for 5min; 40 cycles at 70°C for 40 sec, 65°C or 70°C at 40 sec, 72°C at 50 sec; 1 cycle at 72°C for 10 min and hold of 4°C. An aliquot (3µl of 50 µl) of amplified products were detected by electrophoresis on a 1% agarose gel and visualized by ethidium bromide staining UV illumination and purified with QIAquick PCR Purification Kit (QIAGEN). The amplicons were nucleotide directly sequenced without molecular cloning. Nucleotide sequencing reactions were performed with the ABI BigDye terminator V3.1 Ready Reaction Cycle Sequencing Mixture (Applied Biosystems) according to manufacturer's recommendations. Nucleotide sequence was determined by capillary electrophoresis at Sequencing Facility of Fiocruz-RJ (RPT01A - *Seqüenciamento de DNA* – RJ). Raw sequence data were aligned and edited by using the SeqMan module of LaserGene (DNASTAR Inc.).

The complete genome sequences of both YF viruses samples were elucidated with this approach and they were deposited in the GenBank database under the following accession numbers (waiting for accession numbers – submitted on April 3th, 2017) , for strains ES-504/BRA/2017 and ES-505/BRA/2017, respectively. When we compared these genomes, they displayed 100% of identity. The evolutionary relationships of these

two YFV strains from the ongoing outbreak with the modern YF sequences, mainly from South American genotype I, was established by phylogenetic analysis. Initially, we selected a set of sequences of prM/E junction fragment using Blast tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The 666-bp sequence consists of the last 108 nucleotides of prM gene, including the entire 225 nucleotides of M gene, and the first 333 nucleotides of E gene. Nucleotide sequences were aligned using the CLUSTAL W program (Thompson et al., 1994) with selected YF viral sequences available at the GenBank database. Phylogenetic tree was generated by the Neighbor-joining method (Saitou & Nei, 1987) under a matrix of genetic distances established under the Kimura-two parameter model (Kimura, 1980), by means of the MEGA7 program (Kumar et al., 2016). The robustness of each node was assessed by bootstrap resampling (2,000 replicates) (Felsenstein, 1993). The homologous region (prM/E) of a dengue virus strain available at the GenBank database (PaH881/88; Accession number: AF349753) was used as an outgroup. The Asibi prototype yellow fever strain (Accession number: AY640589) and the vaccine strain 17DD-Brazil (Accession number: DQ100292) were also incorporated in the analysis.

The South American YF sequences in this study formed two major clusters: the South America I and South America II genotypes, supported by 97% and 98% bootstrap values, respectively (Figure 2). The South America I genotype clade is further divided into sub-clades as described by Vasconcelos et al (2004) and de Souza et al (2010). Sequence strains from ES-504/BRA/2017 and ES-505/BRA/2017 belonged to the South America I genotype, and grouped inside 1 E sub-clade in conjunction with other modern strains detected in Brazil (years: 2002, 2004, 2008) and Venezuela (years: 1998, 2005-2007, 2010). The recent Brazilian and Venezuelan strains that were characterized from infections in humans and NHP, also clustered in 1 E sub-clade

(South America I genotype). Auguste and colleagues (2015) suggested that Brazil is the major source of YFV introduction into Venezuela. However, our data suggests that most recent Brazilian YFV strains would have originated from a Venezuelan YFV strain, since oldest strains inside E1 sub-clade were isolated in Venezuela in 1998 (Figure 2). The acquisition in phylogenetic studies of additional complete YF genomes from ancestral and present circulating strains from humans, NHP and mosquitos become evident.

The comparison of the YF virus precursor polyproteins obtained from complete genome sequences with those detected in Brazil and Venezuela since 1980 showed eight unique and semi conservative amino acid changes at C, NS3 and NS5 proteins (Figure 3). They map at the following polyprotein positions: (1) 108 for isoleucine (C protein); (2) 1572 for aspartic acid and 1605 for lysine at NS3 region; (3) 2607 for arginine, 2644 for isoleucine , 2679 for serine , 3149 for alanine and 3215 for serine (3215) at NS5 protein. Interestingly, seven out of eight amino acids changes are located in the two more important proteins of viral replicase complex, the protein NS3 and NS5, and perhaps being associated with some selective advantage in the viral fitness reflecting in its ability to infect vertebrate and/or invertebrate hosts and spreading.

However, it remains to be determined whether these specific amino acid changes are unique to the strains belonging to the ongoing outbreak. Alternatively, they, or at least some of them, could occur in some ancestral sequences that have not been sequenced so far. Hence, there are relatively very few complete genomes from the Americas available at the GenBank database. On the other hand, this matter will be better clarified with the genome elucidation of other circulating YF viruses in the current outbreak from infected mosquito, NHP and human biological samples.

Finally, it is very important to point out that to better understand the molecular epidemiology and evolution of YFV and their potential association with viral spreading and infectivity it is of utmost relevance to determine the ancestor and modern YFV strains.

ACKNOWLEDGEMENTS

To Marta Pereira Santos and Marcelo Quintela Gomes, for technical assistance; to Alessandro Pecego M. Romano (Grupo Técnico de Vigilância de Arboviroses) and Roberta Gomes de Carvalho, (Programa Nacional de Controle da Dengue) Brazilian Ministry of Health, Gilsa Aparecida P. Rodrigues (Secretaria de Saúde do Estado do Espírito Santo), Gilton Luiz Almada (Centro de Informação Estratégica de Vigilância em Saúde-ES) and Roberto da Costa Laterrière Junior (Núcleo Especial de Vigilância Ambiental-ES), for the access to epidemiological data and support for the field work; to Núcleo de Entomologia e Malacologia do Espírito Santo (NEMES), Marilza L. Lange, (Secretaria Municipal de Saúde, Municipality of Domingos Martins), Vigilância Ambiental and Luciano L. Salles (Vigilância em Saúde, Municipality of Ibatiba), for technical support in the field work.

AUTHORS' CONTRIBUTION

MCB and RLO conceived the study; FVSA carried out the collection of biological specimens; RMM, AFB and MGC carried out the viral RNA extraction from biological specimens and the diagnosis by RT-PCR; AACS performed the rapid viral RNA extraction and genome sequencing, MCB, AACS and MMG analyzed the genome sequences, MMG performed phylogenetic analysis; RLO, MCB and MMG prepared

the manuscript. All authors critically read and approved the final version of the manuscript.

References

Auguste AJ, Lemey P, Bergren NA, Giambalvo D, Moncada M, Morón D, Hernandez R, Navarro JC, Weaver SC 2015. Enzootic transmission of yellow fever virus, Venezuela. Emerg Infect Dis. Jan;21(1):99-102.

Bryant JE, Holmes EC, Barrett AD 2007. Out of Africa: a molecular perspective on the introduction of yellow fever virus into the Americas. PLoS Pathog, 3, e75.

COES, Febre Amarela, INFORME – Nº 32/2017. Available from: http://portalarquivos.saude.gov.br/images/pdf/2017/marco/16/COES-FEBRE-AMARELA---INFORME-32---Atualiza----o-em-16mar2017---s-13horas.pdf.

de Souza RP, Foster PG, Sallum MA, Coimbra TL, Maeda AY, Silveira VR, Moreno ES, da Silva FG, Rocco IM, Ferreira IB, Suzuki A, Oshiro FM, Petrella SM, Pereira LE, Katz G, Tengan CH, Siciliano MM, Dos Santos CL 2010. Detection of a new yellow fever virus lineage within the South American genotype I in Brazil. J Med Virol, 82, 175-185.

Felsenstein J 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.

Kimura M 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16:111-120.

Kumar S, Stecher G, and Tamura K 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets.Molecular Biology and Evolution 33:1870-1874.

Monath TP, Vasconcelos PF 2015. Yellow fever. J Clin Virol, 64, 160-173.

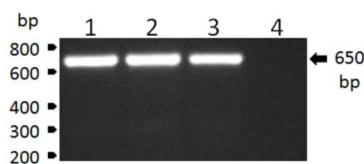
Nunes MR, Palacios G, Cardoso JF, Martins LC, Sousa EC, Jr., de Lima CP, Medeiros DB, Savji N, Desai A, Rodrigues SG, Carvalho VL, Lipkin WI, Vasconcelos PF 2012. Genomic and phylogenetic characterization of Brazilian yellow fever virus strains. J Virol, 86, 13263-13271.

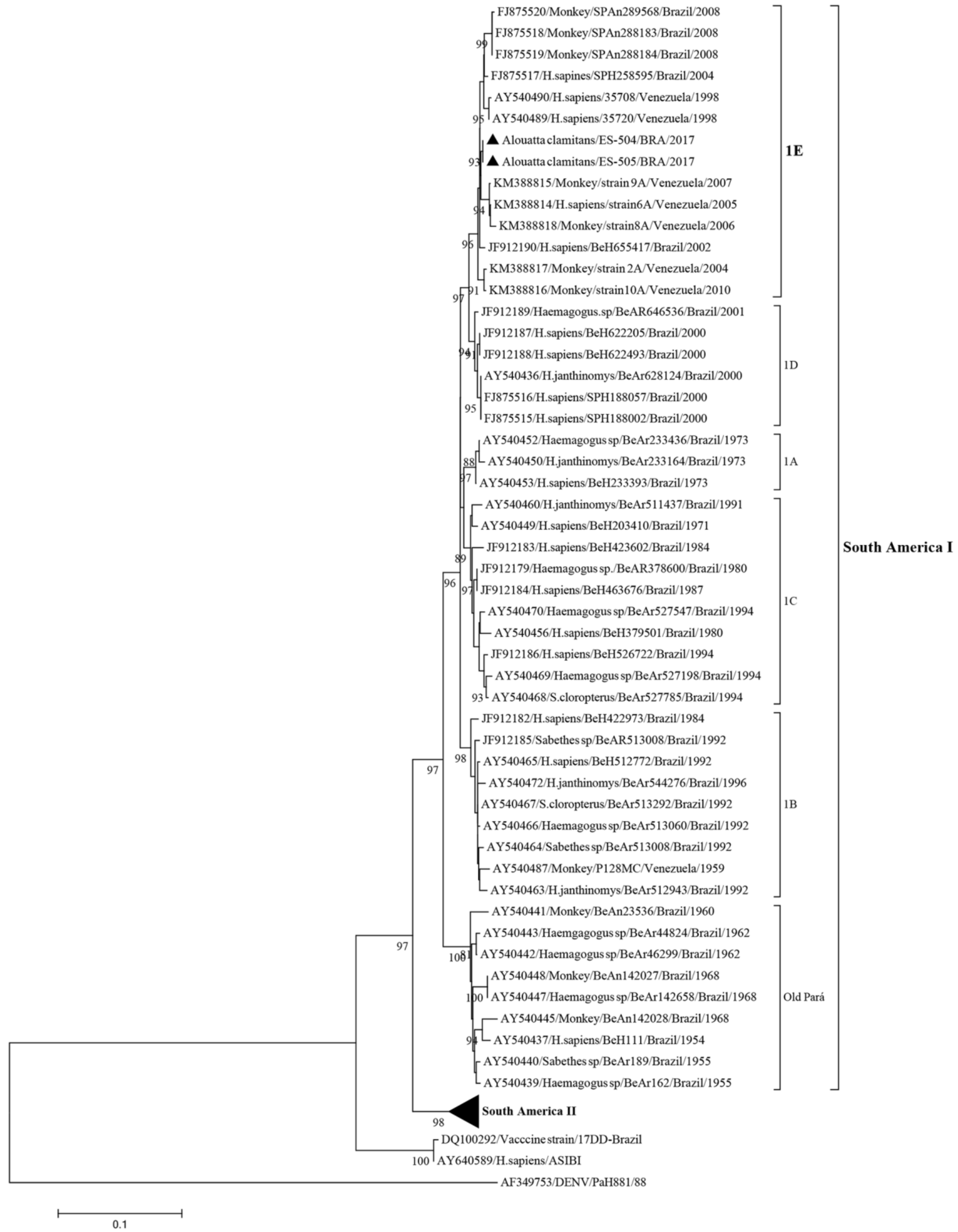
Thompson JD, Higgins DG, and Gibson TJ 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680.

Saitou N. and Nei M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406-425.

Vasconcelos PF, Bryant JE, da Rosa TP, Tesh RB, Rodrigues SG, Barrett AD 2004. Genetic divergence and dispersal of yellow fever virus, Brazil. Emerg Infect Dis, 10, 1578-1584.

Vasconcelos PF, Monath TP 2016. Yellow Fever Remains a Potential Threat to Public Health. Vector Borne Zoonotic Dis, 16, 566-567.




FJ875520/Monkey/SPAn289568/Brazil/2008
FJ875518/Monkey/SPAn288183/Brazil/2008
FJ875519/Monkey/SPAn288184/Brazil/2008
FJ875517/H.sapines/SPH258595/Brazil/2004
AY540490/H.sapiens/35708/Venezuela/1998
AY540489/H.sapiens/35720/Venezuela/1998
Alouatta clamitans/ES-504/BRA/2017
Alouatta clamitans/ES-505/BRA/2017
KM388815/Monkey/strain 9A/Venezuela/2007
KM388814/H.sapiens/strain6A/Venezuela/2005
KM388818/Monkey/strain8A/Venezuela/2006
JF912190/H.sapiens/BeH655417/Brazil/2002
KM388817/Monkey/strain 2A/Venezuela/2004
KM388816/Monkey/strain10A/Venezuela/2010
1E
JF912189/Haemagogus.sp/BeAR646536/Brazil/2001
JF912187/H.sapiens/BeH622205/Brazil/2000
JF912188/H.sapiens/BeH622493/Brazil/2000
AY540436/H.janthinomys/BeAr628124/Brazil/2000
FJ875516/H.sapiens/SPH188057/Brazil/2000
FJ875515/H.sapiens/SPH188002/Brazil/2000
1D
AY540452/Haemagogus sp/BeAr233436/Brazil/1973
AY540450/H.janthinomys/BeAr233164/Brazil/1973
AY540453/H.sapiens/BeH233393/Brazil/1973
1A
AY540460/H.janthinomys/BeAr511437/Brazil/1991
AY540449/H.sapiens/BeH203410/Brazil/1971
JF912183/H.sapiens/BeH423602/Brazil/1984
JF912179/Haemagogus sp./BeAR378600/Brazil/1980
JF912184/H.sapiens/BeH463676/Brazil/1987
AY540470/Haemagogus sp/BeAr527547/Brazil/1994
AY540456/H.sapiens/BeH379501/Brazil/1980
JF912186/H.sapiens/BeH526722/Brazil/1994
AY540469/Haemagogus sp/BeAr527198/Brazil/1994
AY540468/S.cloropterus/BeAr527785/Brazil/1994
1C
South America I
JF912182/H.sapiens/BeH422973/Brazil/1984
JF912185/Sabethes sp/BeAR513008/Brazil/1992
AY540465/H.sapiens/BeH512772/Brazil/1992
AY540472/H.janthinomys/BeAr544276/Brazil/1996
AY540467/S.cloropterus/BeAr513292/Brazil/1992
AY540466/Haemagogus sp/BeAr513060/Brazil/1992
AY540464/Sabethes sp/BeAr513008/Brazil/1992
AY540487/Monkey/P128MC/Venezuela/1959
AY540463/H.janthinomys/BeAr512943/Brazil/1992
1B
AY540441/Monkey/BeAn23536/Brazil/1960
AY540443/Haemgagogus sp/BeAr44824/Brazil/1962
AY540442/Haemagogus sp/BeAr46299/Brazil/1962
AY540448/Monkey/BeAn142027/Brazil/1968
AY540447/Haemagogus sp/BeAr142658/Brazil/1968
AY540445/Monkey/BeAn142028/Brazil/1968
AY540437/H.sapiens/BeH111/Brazil/1954
AY540440/Sabethes sp/BeAr189/Brazil/1955
AY540439/Haemagogus sp/BeAr162/Brazil/1955
Old Pará
South America II
DQ100292/Vacccine strain/17DD-Brazil
AY640589/H.sapiens/ASIBI
AF349753/DENV/PaH881/88
0.1

Protein: C (Anchor C) | prM (M) | E | NS1 | NS2A | NS2B | NS3 | NS4A | NS4B | NS5

Domain: Data

Amino acid position

Clade	Protein id / Strain / Country / Year	Host	9	79	82	105	**108**	119	127	128	163	258	405	432	462	483	787	808	881	1064	1110	1116	1133	1136	1146	1150	1178	1185
	ES-505/BRA/2017	Monkey	K	V	R	V	I	V	R	N	P	A	T	A	R	M	G	N	P	V	P	K	I	I	A	V	I	T
	AFH35044/BeH655417/Brazil/2002	Human	K	V	R	V	V	V	R	N	P	A	T	A	R	M	G	N	P	V	P	K	V	I	A	V	V	T
	AIZ07887/strain10A/Venezuela/2010	Monkey	K	V	R	V	V	V	R	N	P	A	T	A	R	M	G	N	P	V	P	K	I	I	A	V	I	T
1E	AIZ07885/strain6A/Venezuela/2005	Human	K	V	R	V	V	V	R	N	P	A	T	A	R	M	G	N	P	V	P	K	I	I	A	I	I	T
	AIZ07889/strain8A/Venezuela/2006	Monkey	K	V	R	V	V	V	R	N	P	A	T	A	R	M	G	N	P	V	P	K	I	I	A	I	I	T
	AIZ07888/strain 2A/Venezuela/2004	Monkey	K	V	R	V	V	A	R	N	P	T	T	A	R	M	G	N	P	A	P	K	I	I	A	V	I	T
	AIZ07886/strain 9A/Venezuela/2007	Monkey	K	V	R	V	V	V	R	N	P	A	T	A	R	M	G	N	P	V	P	K	I	I	A	I	I	I
	AFH35043/BeAR646536/Brazil/2001	Mosquito	K	A	K	V	V	V	R	N	P	A	T	V	R	M	G	N	P	V	S	K	I	V	A	V	I	T
1D	AFH35042/BeH622493/Brazil/2000	Human	K	A	K	V	V	V	R	N	P	A	T	A	R	M	G	N	P	V	S	K	I	V	A	V	I	T
	AFH35041/BeH622205/Brazil/2000	Human	K	A	K	V	V	V	R	N	S	A	T	A	R	M	G	N	P	V	S	K	I	V	A	V	I	T
1B	AFH35036/BeH422973/Brazil/1984	Human	K	A	K	V	V	V	K	N	P	A	A	A	R	M	G	N	P	V	P	K	I	I	A	V	I	T
	AFH35039/BeAR513008/Brazil/1992	Mosquito	K	A	K	V	V	V	K	N	P	A	T	A	R	M	G	N	P	V	P	R	I	V	A	V	I	T
	AFH35037/BeH423602/Brazil/1984	Human	K	A	K	V	V	V	R	N	P	A	T	A	K	M	G	S	P	V	P	K	I	I	A	V	I	T
1C	AFH35033/BeAR378600/Brazil/1980	Mosquito	K	A	K	A	V	V	R	N	P	A	T	A	K	M	S	S	P	V	P	K	I	I	A	V	I	T
	AFH35040/BeH526722/Brazil/1994	Human	R	A	K	V	V	V	R	S	P	A	T	A	K	M	G	S	P	V	P	K	I	I	T	V	I	T
	AFH35038/BeH463676/Brazil/1987	Human	K	A	K	A	V	V	R	N	P	A	T	A	K	R	G	S	S	V	P	K	I	I	A	V	I	T

Clade	Protein id / Strain / Country / Year	Host	1197	1249	1277	1291	1308	1313	1408	1457	1477	1482	1550	**1572**	**1605**	1669	1679	1701	1811	1835	1927	2006	2045	2137	2276	2311	2408	2502
	ES-505/BRA/2017	Monkey	N	M	V	A	V	S	R	A	F	A	K	D	K	P	I	A	S	T	R	V	T	L	I	H	N	K
	AFH35044/BeH655417/Brazil/2002	Human	N	M	V	A	V	S	R	A	F	A	K	E	R	P	I	A	N	T	R	V	T	L	I	H	N	K
	AIZ07887/strain10A/Venezuela/2010	Monkey	N	M	V	A	V	S	R	A	F	A	K	E	R	P	I	A	S	T	R	V	T	L	I	H	N	K
1E	AIZ07885/strain6A/Venezuela/2005	Human	N	M	V	A	V	T	R	A	F	T	K	E	R	P	I	A	S	T	R	V	T	L	I	H	N	K
	AIZ07889/strain8A/Venezuela/2006	Monkey	N	M	V	A	V	T	R	A	F	T	K	E	R	P	I	A	S	K	R	V	T	L	I	L	S	N
	AIZ07888/strain 2A/Venezuela/2004	Monkey	N	M	V	A	V	S	R	A	F	A	K	E	R	P	I	A	S	T	R	V	T	L	I	H	N	K
	AIZ07886/strain 9A/Venezuela/2007	Monkey	S	M	V	A	V	T	R	A	F	T	K	E	R	P	V	A	S	T	R	V	T	L	I	H	S	K
	AFH35043/BeAR646536/Brazil/2001	Mosquito	N	M	V	A	V	S	R	A	F	A	K	E	R	S	I	A	S	T	R	V	T	L	I	H	N	K
1D	AFH35042/BeH622493/Brazil/2000	Human	N	M	V	A	V	S	R	A	F	A	K	E	R	P	I	A	S	T	R	V	T	L	I	H	N	K
	AFH35041/BeH622205/Brazil/2000	Human	N	M	V	A	V	S	R	A	F	A	K	E	R	P	I	T	S	T	R	V	T	L	I	H	N	K
1B	AFH35036/BeH422973/Brazil/1984	Human	N	T	V	A	V	S	R	A	F	A	R	E	R	P	I	A	S	T	R	V	T	L	I	H	N	K
	AFH35039/BeAR513008/Brazil/1992	Mosquito	N	T	A	A	V	S	R	A	F	A	K	E	R	P	I	A	S	T	K	V	M	L	T	H	N	K
	AFH35037/BeH423602/Brazil/1984	Human	N	M	V	A	I	S	K	A	L	A	K	E	R	P	I	A	S	T	R	V	T	F	I	H	N	N
1C	AFH35033/BeAR378600/Brazil/1980	Mosquito	N	M	V	A	I	S	K	A	L	A	K	E	R	P	I	A	S	T	R	V	T	F	I	H	N	N
	AFH35040/BeH526722/Brazil/1994	Human	S	M	V	V	I	S	K	A	L	A	K	E	R	P	I	A	S	T	R	A	T	F	I	H	N	K
	AFH35038/BeH463676/Brazil/1987	Human	N	M	V	A	I	S	K	T	L	A	K	E	R	P	I	A	S	T	R	V	T	F	I	H	N	N

Clade	Protein id / Strain / Country / Year	Host	2503	2601	**2607**	2614	2638	2640	**2644**	2647	2648	**2679**	2736	2803	2903	2907	3031	**3149**	3152	3158	**3215**	3234	3245	3310	3341	3346	3363	3375
	ES-505/BRA/2017	Monkey	T	R	R	E	V	R	I	D	T	S	T	S	K	I	A	A	K	A	S	I	V	K	K	I	M	L
	AFH35044/BeH655417/Brazil/2002	Human	T	R	K	E	V	R	V	D	T	G	T	N	K	I	A	V	K	A	N	I	V	K	K	I	M	L
	AIZ07887/strain10A/Venezuela/2010	Monkey	T	R	K	E	V	R	V	D	T	G	T	N	K	I	T	V	K	T	N	I	V	K	K	I	M	L
1E	AIZ07885/strain6A/Venezuela/2005	Human	I	R	K	E	V	R	V	D	T	G	T	N	K	I	A	V	K	A	N	I	V	K	K	I	M	L
	AIZ07889/strain8A/Venezuela/2006	Monkey	T	R	K	E	V	R	V	D	T	G	T	S	K	I	A	V	K	A	N	I	V	K	K	I	M	L
	AIZ07888/strain 2A/Venezuela/2004	Monkey	T	R	K	K	V	R	V	D	T	G	T	N	K	I	A	V	K	A	N	I	V	K	K	I	M	L
	AIZ07886/strain 9A/Venezuela/2007	Monkey	T	R	K	E	V	R	V	D	T	G	T	N	K	I	A	V	K	A	N	I	V	K	K	I	M	L
	AFH35043/BeAR646536/Brazil/2001	Mosquito	T	R	K	E	V	P	V	V	I	G	A	D	K	I	A	V	K	A	N	I	V	R	K	I	M	L
1D	AFH35042/BeH622493/Brazil/2000	Human	T	R	K	E	V	R	V	D	T	G	A	D	K	I	A	V	K	A	N	I	V	R	K	I	M	L
	AFH35041/BeH622205/Brazil/2000	Human	T	R	K	E	V	R	V	D	T	G	A	D	K	I	A	V	K	A	N	I	V	R	K	I	M	L
1B	AFH35036/BeH422973/Brazil/1984	Human	T	R	K	E	V	R	V	D	T	G	A	N	K	I	A	V	K	A	N	I	V	K	K	I	I	L
	AFH35039/BeAR513008/Brazil/1992	Mosquito	T	R	K	E	V	R	V	D	T	G	A	N	K	M	A	V	K	A	N	V	I	K	K	I	M	F
	AFH35037/BeH423602/Brazil/1984	Human	T	R	K	E	I	R	V	D	T	G	T	N	R	I	A	V	K	A	N	I	V	K	R	I	M	L
1C	AFH35033/BeAR378600/Brazil/1980	Mosquito	T	K	K	E	I	R	V	D	T	G	T	N	K	I	A	V	R	A	N	I	V	K	R	I	M	L
	AFH35040/BeH526722/Brazil/1994	Human	T	K	K	E	I	R	V	D	T	G	T	N	K	I	A	V	K	A	N	I	V	K	R	V	M	L
	AFH35038/BeH463676/Brazil/1987	Human	T	K	K	E	I	R	V	D	T	G	T	N	K	I	A	V	R	A	N	I	V	K	R	I	M	L

Table 1. Primers used for amplification and sequencing of the whole genome of Brazilian yellow fever viruses described in the current study.

Primer	Genome Position	Sequence, 5'→3'	Amplicon size (bp)
YF1F	1-27	AGTAAATCCTGTGTGCTAATTGAGG TG	1400
YF1R	1380-1400	GCTGTGCCCTAATGACATACT	1400
YF2F	985-1006	AGGAATAACCGACAGGGATTTC	1339
YF2R	2304-2323	GCTCAAGCCACCAAACAATC	1339
YF3F	2144-2163	AAGGAAGGCAGCTCAATAGG	1444
YF3R	3566-3587	GTCCCTGTCTCTTCCTCAATAC	1444
YF4F	3426-3446	GGAGTGATGGTTGCTGGTATT	1449
YF4R	4853-4874	CTGCTATCAACTGAACCTCCTC	1449
YF5F	4506-4526	TCCACCCATTTGCACTCTTAC	586
YF5R	5070-5091	TTCACCTCAGTTTGGGATATGG	586
YF6F	4821-4842	CATGGAAGTTGGAGGGTAGATG	1358
YF6R	6156-6178	GACTCTCCTCTGGTCATCTCTTA	1358
YF7F	5840-5862	GTGGAGAGAGTGTTGGATTGTAG	916
YF7R	6736-6755	AAGAGATGTGGGTGGGTTTG	916
YF8F	6089-6108	GGGATGGTTGCTCCACTTTA	1440
YF8R	7507-7528	GAGGCTGGTATTTCCCTCTATG	1440
YF9F	7426-7445	CTTGGCCTCTGTTGCTATGT	1235
YF9R	8640-8660	TTCTCGTGACCTCCTCTATCC	1235
YF10F	8533-8554	CCCTTACAGGACTTGGCATTAT	1257
YF10R	9769-9789	TGAAAGTGGTGAGAGCAGAAG	1257
YF11F	9249-9268	GATGGGACACACGCATAACA	1441
YF11R	10670-10689	GGGCTGACATCCCACTATTT	1441
YF12F	10312-10340	GTACTCTGTGGATGCTGATCTGCAG CCCG	697
YF12R	10965-11008	AGTGGTTTTGTGTTTGTCATCCAAA GGTCTGCTTATTCTTGAGC	697