

## **Supplementary information**

### **Phylogenetics of Yellow Fever Virus in the Americas: new insights into the origin of the 2017 Brazilian outbreak**

Daiana Mir<sup>1</sup>, Edson Delatorre<sup>\*1</sup>, Myrna Bonaldo<sup>2</sup>, Ricardo Lourenço-de-Oliveira<sup>3</sup>, Ana Carolina Vicente<sup>4</sup> and Gonzalo Bello<sup>1</sup>

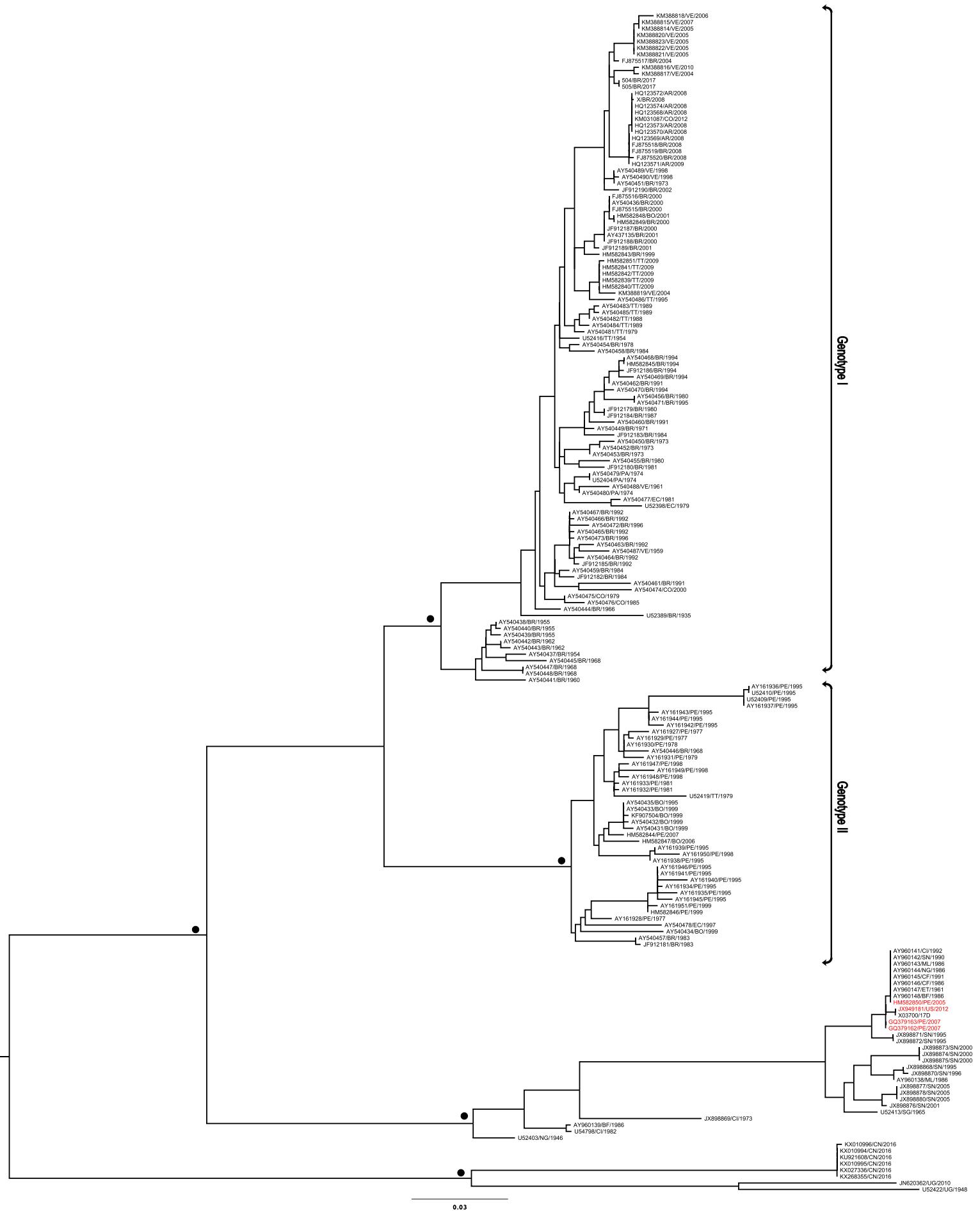
1. Laboratório de AIDS e Imunologia Molecular, Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, Brazil.

2. Laboratório de Biologia Molecular de Flavivírus, Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, Brazil.

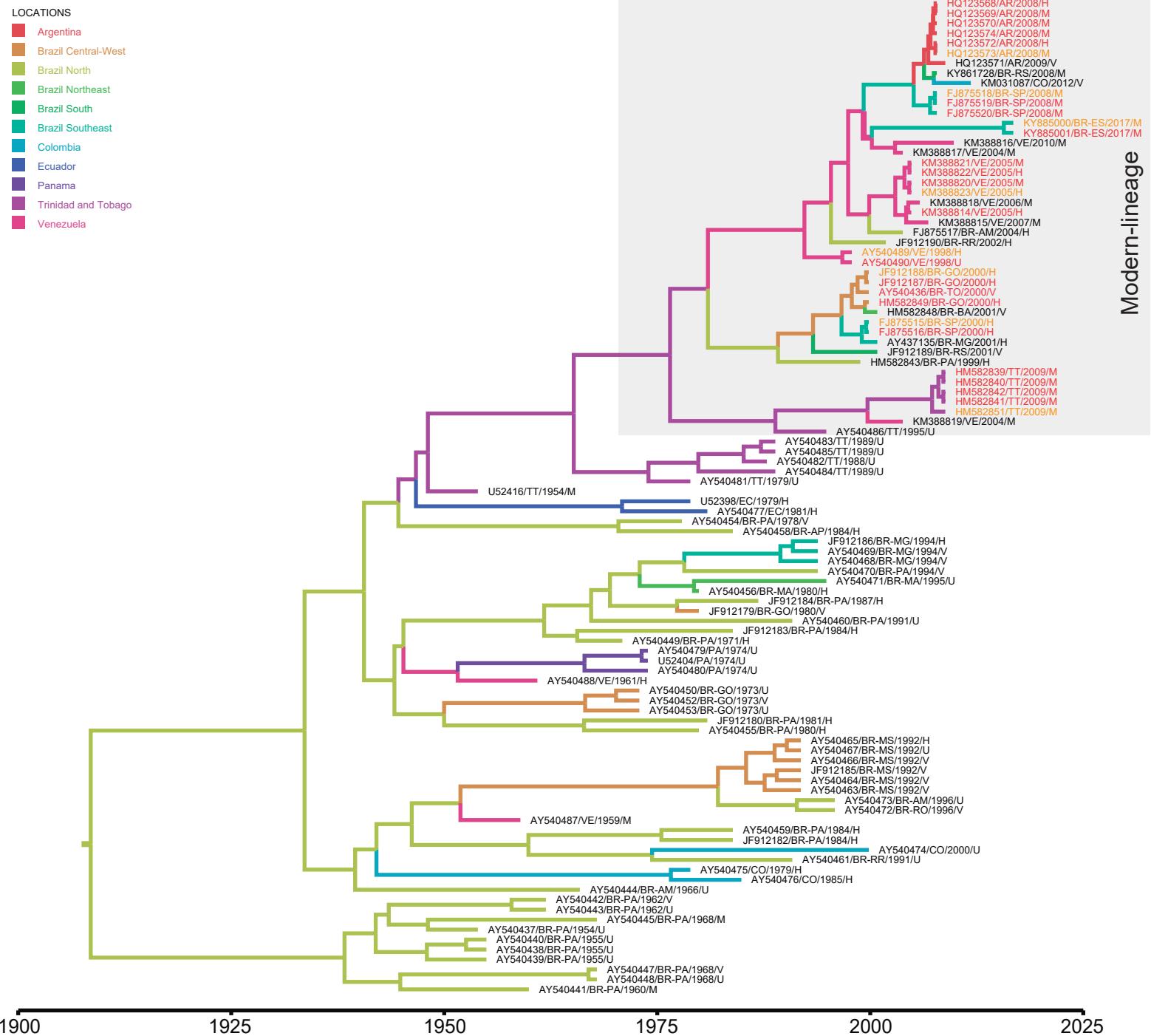
3. Laboratório de Mosquitos Transmissores de Hematozoários, Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, Brazil.

4. Laboratório de Genética Molecular de Microrganismos, Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, Brazil.

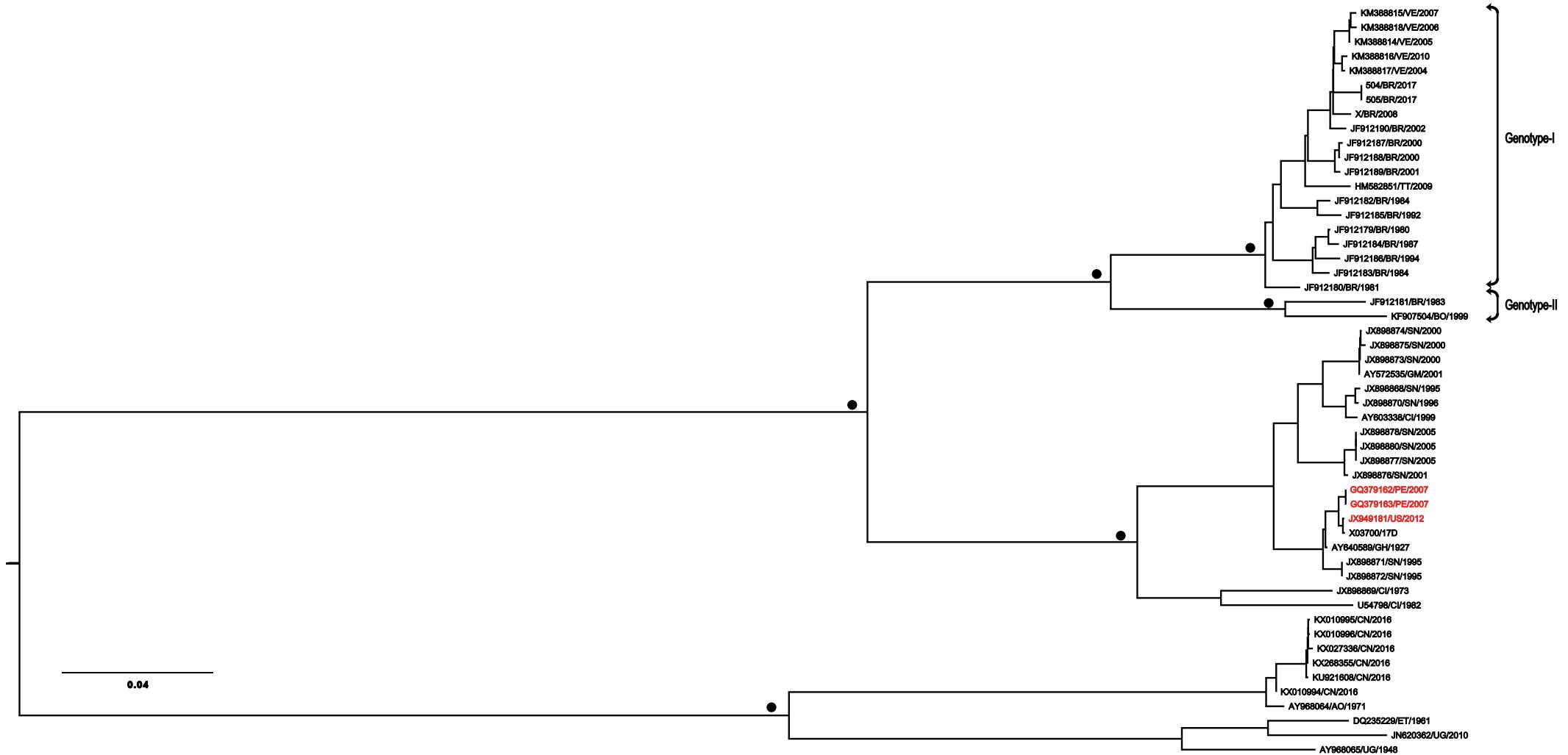
\* Corresponding author: edsonod@ioc.fiocruz.br/delatorre.ioc@gmail.com



**Supplementary Figure S1. Maximum likelihood phylogeny of YFV prM/E gene sequences.** The aLRT support value of each indicated clade (dark full circle) was >0.8. American sequences that grouped with vaccine strain 17D are highlighted in red. The branch lengths are drawn to scale with bar at the bottom indicating nucleotide substitutions per site.



**Supplementary Figure S2. Time-scaled Bayesian MCC phylogeographic tree of the YFV genotype I prM/E gene sequences.** Branches are colored according to the most probable location state of their descendant nodes as indicated at the legend (top left). Tip labels of the modern-lineage (gray shaded area) sequences are colored indicating those epidemiologically related and sampled at the same YFV outbreak (red) and the one that was chosen as the representative of the location and year (orange). All horizontal branch lengths are drawn to a scale of years. The tree is automatically rooted under the assumption of a relaxed molecular clock.



**Supplementary Figure S3. Maximum likelihood phylogeny of YFV complete coding region sequences.** The aLRT support value of each indicated clade (dark full circle) was = 1.0. American sequences that grouped with vaccine strain 17D are highlighted in red. The branch lengths are drawn to scale with bar at the bottom indicating nucleotide substitutions per site.

**Supplementary Table S1.** Best fit demographic model for YFV genotype II prM/E dataset.

Clade	Coalescent model	PS Log ML	Models compared	Log BF	SS Log ML	Models compared	Log BF
	Exponential	-2472.9	-	-	-2473.0	-	-
	BSKY	-2475.2	Exponential/BSKY	2.3	-2475.5	Exponential/BSKY	2.5
	Constant	-2480.6	Exponential/Constant BSKY/Constant	7.7 5.4	-2480.7	Exponential/Constant BSKY/Constant	7.7 5.2
YFV-GII	Expansion	-2480.3	Exponential/Expansion BSKY/Expansion	7.4 5.1	-2480.5	Exponential/Expansion BSKY/Expansion	7.5 5.0
	Logistic	-2481.1	Exponential/Logistic BSKY/Logistic	8.2 5.9	-2481.5	Exponential/Logistic BSKY/Logistic	8.5 6.0

Log marginal likelihood (ML) estimates for the Bayesian skyline, logistic, exponential and expansion growth demographic models obtained using the path sampling (PS) and stepping-stone sampling (SS) methods. The Log Bayes factor (BF) is the difference of the Log ML between of alternative (H1) and null (H0) models (H1/H0). Log BFs > 3 indicates that model H1 is more strongly supported by the data than model H0.





**Supplementary Table S3** - Amino acid mutations with respect to ancestor N3 found in the modern subclade 1E sequences

Name	Type	CDS Position	CDS Codon Number	Polymorphism Type	Mature peptide	Change	Codon position	Codon Change	Amino Acid Change	Variant Sequences (Accession number)
M_1	Nonsynonymous	322	108	transition	aC	G->A	1	GTA->ATA	V->I	KY885000; KY885001
M_2	Nonsynonymous	356	119	transition	aC	T->C	2	GTT->GCT	V->A	KM388817
M_3	Nonsynonymous	772	258	transition	M	G->A	1	GCC->ACC	A->T	KM388817
M_4	Nonsynonymous	3185	1062	transition	NS1	G->A	2	AGT->AAT	S->N	KY861728
M_5	Nonsynonymous	3191	1064	transition	NS1	T->C	2	GTT->GCT	V->A	KM388817
M_6	Nonsynonymous	3328	1110	transition	NS1	C->T	1	CCT->TCT	P->S	KY861728
M_7	Nonsynonymous	3397	1133	transition	NS2a	A->G	1	ATC->GTC	I->V	JF912190
M_8	Nonsynonymous	3448	1150	transition	NS2a	G->A	1	GTA->ATA	V->I	KM388814; KM388815; KM388818
M_9	Nonsynonymous	3532	1178	transition	NS2a	A->G	1	ATC->GTC	I->V	JF912190
M_10	Nonsynonymous	3554	1185	transition	NS2a	C->T	2	ACA->ATA	T->I	KM388815
M_11	Nonsynonymous	3590	1197	transition	NS2a	A->G	2	AAT->AGT	N->S	KM388815
M_12	Nonsynonymous	3937	1313	transversion	NS2a	T->A	1	TCA->ACA	S->T	KM388814; KM388815; KM388818
M_13	Nonsynonymous	4444	1482	transition	NS2b	G->A	1	GCC->ACC	A->T	KM388814; KM388815; KM388818
M_14	Nonsynonymous	4716	1572	transversion	NS3	G->T	3	GAG->GAT	E->D	KY885000; KY885001
M_15	Nonsynonymous	4773	1591	transversion	NS3	C->A	3	AAC->AAA	N->K	KM388818
M_16	Nonsynonymous	4814	1605	transition	NS3	G->A	2	AGG->AAG	R->K	KY885000; KY885001
M_17	Nonsynonymous	4997	1666	transition	NS3	A->G	2	CAG->CGG	Q->R	KY861728
M_18	Nonsynonymous	5035	1679	transition	NS3	A->G	1	ATT->GTT	I->V	KM388815
M_19	Nonsynonymous	5432	1811	transition	NS3	G->A	2	AGT->AAT	S->N	JF912190
M_20	Nonsynonymous	5504	1835	transversion	NS3	C->A	2	ACA->AAA	T->K	KM388818
M_21	Nonsynonymous	6932	2311	transversion	NS4b	A->T	2	CAC->CTC	H->L	KM388818
M_22	Nonsynonymous	7223	2408	transition	NS4b	A->G	2	AAT->AGT	N->S	KM388815; KM388818
M_23	Nonsynonymous	7506	2502	transversion	NS4b	G->T	3	AAG->AAT	K->N	KM388818
M_24	Nonsynonymous	7508	2503	transition	NS4b	C->T	2	ACT->ATT	T->I	KM388814
M_25	Nonsynonymous	7802	2601	transition	NS5	G->A	2	AGA->AAA	R->K	KY861728
M_26	Nonsynonymous	7820	2607	transition	NS5	A->G	2	AAG->AGG	K->R	KY885000; KY885001
M_27	Nonsynonymous	7840	2614	transition	NS5	G->A	1	GAA->AAA	E->K	KM388817
M_28	Nonsynonymous	7930	2644	substitution	NS5	GTG->ATA	1	GTG->ATA	V->I	KY885000; KY885001
M_29	Nonsynonymous	8035	2679	transition	NS5	G->A	1	GGC->AGC	G->S	KY885000; KY885001
M_30	Nonsynonymous	8408	2803	transition	NS5	A->G	2	AAT->AGT	N->S	KY885000; KY885001; KM388818
M_31	Nonsynonymous	9091	3031	transition	NS5	G->A	1	GCA->ACA	A->T	KM388816
M_32	Nonsynonymous	9446	3149	transition	NS5	T->C	2	GTT->GCT	V->A	KY885000; KY885001
M_33	Nonsynonymous	9472	3158	transition	NS5	G->A	1	GCT->ACT	A->T	KM388816
M_34	Nonsynonymous	9644	3215	transition	NS5	A->G	2	AAT->AGT	N->S	KY885000; KY885001
M_35	Nonsynonymous	10033	3345	transition	NS5	G->A	1	GAC->AAC	D->N	KY861728

**Supplementary Table S4** - List of the GenBank accession numbers of the YFV sequences used in this study

Country	Region	Genbank accession number	Collection date
Argentina	-	HQ123568	2008
	-	HQ123569	2008
	-	HQ123570	2008
	-	HQ123572	2008
	-	HQ123573	2008
	-	HQ123574	2008
	-	HQ123571	2009
Bolivia	-	AY540435	1995
	-	AY540431	1999
	-	AY540432	1999
	-	AY540433	1999
	-	AY540434	1999
	-	KF907504	1999
	-	HM582847	2006
Brazil	BR-AM	AY540444	1966
		AY540473	1996
		FJ875517	2004
	BR-AP	AY540458	1984
	BR-BA	HM582848	2001
	BR-ES	KY885000	2017
		KY885001	2017
	BR-GO	AY540450	1973
		AY540452	1973
		AY540453	1973
		JF912179	1980
		HM582849	2000
		JF912187	2000
		JF912188	2000
	BR-MA	AY540456	1980
		AY540471	1995
	BR-MG	AY540468	1994
		AY540469	1994
		JF912186	1994
		AY437135	2001
	BR-MS	AY540463	1992
		AY540464	1992
		AY540465	1992
		AY540466	1992
		AY540467	1992
		JF912185	1992
	BR-PA	AY540437	1954
		AY540438	1955
		AY540439	1955
		AY540440	1955
		AY540441	1960

		AY540442	1962
		AY540443	1962
		AY540445	1968
		AY540446	1968
		AY540447	1968
		AY540448	1968
		AY540449	1971
		AY540454	1978
		AY540455	1980
		JF912180	1981
		AY540459	1984
		JF912182	1984
		JF912183	1984
		JF912184	1987
		AY540460	1991
		AY540470	1994
		HM582843	1999
	BR-RO	AY540457	1983
		JF912181	1983
		AY540472	1996
	BR-RR	AY540461	1991
		JF912190	2002
	BR-RS	JF912189	2001
		KY861728	2008
	BR-SP	FJ875515	2000
		FJ875516	2000
		FJ875518	2008
		FJ875519	2008
		FJ875520	2008
	BR-TO	AY540436	2000
Colombia	-	AY540475	1979
	-	AY540476	1985
	-	AY540474	2000
	-	KM031087	2012
Ecuador	-	U52398	1979
	-	AY540477	1981
	-	AY540478	1997
Panama	-	AY540479	1974
	-	AY540480	1974
	-	U52404	1974
Peru	-	AY161927	1977
	-	AY161928	1977
	-	AY161929	1977
	-	AY161930	1978
	-	AY161931	1979
	-	AY161932	1981
	-	AY161933	1981
	-	AY161934	1995
	-	AY161935	1995
	-	AY161938	1995

	-	AY161939	1995
	-	AY161940	1995
	-	AY161941	1995
	-	AY161942	1995
	-	AY161943	1995
	-	AY161944	1995
	-	AY161945	1995
	-	AY161946	1995
	-	AY161947	1998
	-	AY161948	1998
	-	AY161949	1998
	-	AY161950	1998
	-	AY161951	1999
	-	HM582846	1999
	-	HM582844	2007
Trinidad and Tobago	-	U52416	1954
	-	AY540481	1979
	-	U52419	1979
	-	AY540482	1988
	-	AY540483	1989
	-	AY540484	1989
	-	AY540485	1989
	-	AY540486	1995
	-	HM582839	2009
	-	HM582840	2009
	-	HM582841	2009
	-	HM582842	2009
	-	HM582851	2009
	-	AY540487	1959
	-	AY540488	1961
Venezuela	-	AY540489	1998
	-	AY540490	1998
	-	KM388817	2004
	-	KM388819	2004
	-	KM388814	2005
	-	KM388820	2005
	-	KM388821	2005
	-	KM388822	2005
	-	KM388823	2005
	-	KM388818	2006
	-	KM388815	2007
	-	KM388816	2010