ELSEVIER

PII: S0264-410X(97)00278-8

Vaccine, Vol. 16, No. 9/10, pp. 1024–1028, 1998

Printed in Great Britain
0264-410X/98 $19+0.00

Genetic variability among yellow fever virus 17D substrains

Ricardo Galler*†‡, Paulo R. Post* Claudia N.D. Santos* and Idevaldo I. Ferreira*†

The complete nucleotide sequence of the genome from two yellow fever (YF) virus strains, 17DD and 17D-213 was determined. Comparison of these sequences with those of other YF viruses, including the parental virulent Asibi strain, allowed the identification of 48 nucleotide sequence differences which are 17D strain-specific and potentially related to viral attenuation. Another 43 nucleotide sequence differences were not common to all 17D substrains and are therefore substrain specific. Of the 21 changes between 17DD and Asibi 15 only five led to amino acid substitutions whereas 13 substrain differences common to all 17D-204 substrains produced six amino acid substitutions. Since the exact passage histories of these viruses is known it was possible to calculate, for each strain, the number of accumulated changes per passage. Based on these data the 17DD strain was the most genetically stable virus. © 1998 Elsevier Science Ltd. All rights reserved

Keywords: yellow fever virus; flaviviruses; genetic variation

The *Flavivirus* genus within the *Flaviviridae* is composed of about 70 viruses. Yellow fever (YF) virus is the prototype virus with a genome of 10.862 nucleotides with a 5′ CAP structure and a nonpolyadenylated 3′ end encoding a polyprotein of 3411 amino acids which is cleaved by proteolytic processing to give rise to 11 viral polypeptides. Nucleotide sequence analyses of flavivirus genomes have led to new insights into genome structure and replication[1].

The isolation of YF virus[2] allowed the later development of an attenuated virus (17D)[3] which has been used for over 50 years for human vaccination. Comparison of the genomic sequences of the Asibi strain with that of the 17D-204 virus[4] revealed 67 and 31 nucleotide/amino acid sequence changes, respectively, scattered along the genome. The availability of an animal system that reflects human infection[5,6], and YF infectious cDNA[7] with an established phenotype[8] should allow the identification of genetic determinants of viral virulence and attenuation. One approach to determine the changes responsible for attenuation is to use recombinant DNA methodology to produce 17D/Asibi recombinants and test their resulting phenotypes. Given the large number of differences between these strains it is of interest to try to limit the number of mutations to be tested.

Vaccines currently in use are derived from two distinct substrains (17D-204 and 17DD). If standardization of YF substrain use for vaccine production is desirable then it is important to analyse the genetic stability of each of these viruses as one parameter to be considered upon choosing one substrain. Previous genomic variability analysis using oligonucleotide fingerprinting suggested a high degree of genetic similarity between vaccines produced worldwide with an estimated sequence homology of 98–100%[9]. However, genetic changes were detected and may have occurred within 1–2 passages possibly due to the selection of virion subpopulations or to point mutations. It was suggested that the 17DD-derived vaccines appeared to have more genetic variability than those derived from the 17D-204 substrains. Examination of the extent of genetic variability among different substrains of well characterized YF vaccine viruses may provide clues to those changes which are most likely to be important for attenuation. We have previously found that from the 67 nucleotide sequence changes found to be 17D-specific, as originally proposed[4], only 48 have been confirmed to be common to the three YF 17D virus strains (17D-204, 17D-213 and 17DD) and are, therefore potentially related to viral attenuation[9]. Here we describe those nucleotide sequence changes which are not related to attenuation but instead reflect different histories of independently passaged YF vaccine virus 17D substrains from the original 17D. These results suggest that 17DD is indeed the most genetically stable substrain.

*Departamento de Bioquimica e Biologia Molecular, Fundação Oswaldo Cruz, Rio de Janeiro, RJ Brazil 21045-900. †Instituto de Tecnologia em Imunobiológicos, Fundação Oswaldo Cruz, Rio de Janeiro, RJ Brazil 21045-900. ‡Author to whom all correspondence should be addressed. Tel. 55 21 987 6880; Fax: 55 21 590 3495; e-mail: rgaller@gene.dbbm.fiocruz.br. (Received 28 May 1997; revised version received 20 October 1997; accepted 20 October 1997)

RESULTS

The passage histories of YF virus 17DD and 17D-213 strains are shown in *Figure 1*. Briefly, 17DD was derived at passage 195, with further passages in tissue culture until passage 243 when it was passaged in embryonated eggs until its current passage, 284. Virus recovered from embryo homogenate has been administered to humans for more than half a century. The 17D-213 strain, currently at passage 240, is a derivative of 17D-204. The 17D-213 strain is a leucosis-free virus that has been used for the propagation and preparation of experimental vaccine lots which have been tested for monkey neurovirulence by several laboratories (R.S. Marchevsky, personal communication). The 17D-213 virus used in this study corresponds to the supernatant of primary cultures of chicken embryo fibroblasts (passage 240) which displayed an attenuated phenotype in monkeys. Aiming at the definition of genetic determinants of YF virus attenuation and the extent of genetic variability among strains known to be attenuated, we have determined the complete nucleotide sequence of the genome from two YF vaccine strains (17DD and 17D-213) with known passage histories and attenuated phenotype. We used direct sequencing of RT–PCR amplified RNA since it allows the determination of the majority sequence for the vaccine populations. Thus, the sequences reported for 17DD and 17D-213 should be representative of the viral populations for which the attenuated phenotype has been established[10].

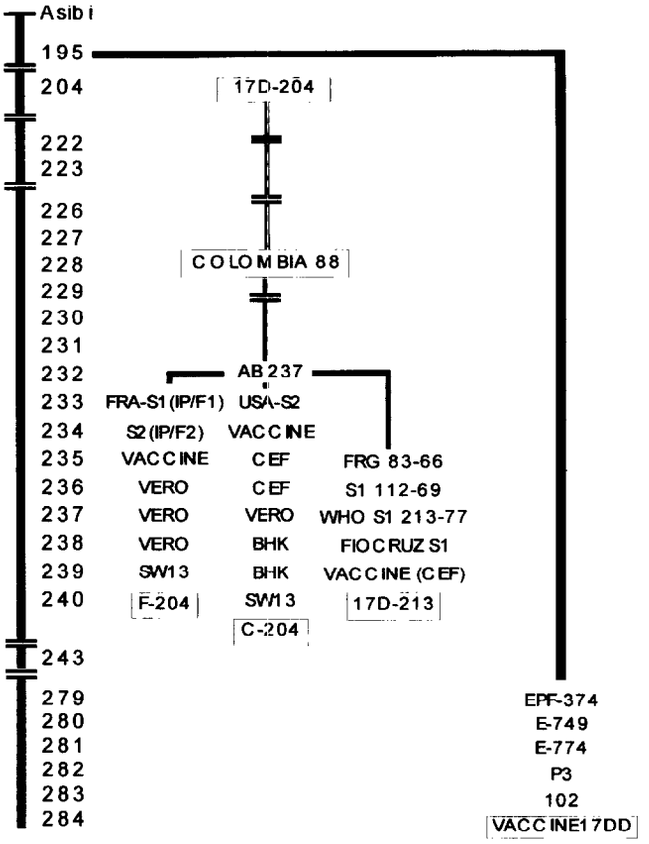

Figure 1 Passage history of the original YF Asibi strain and derivation of YF 17D vaccine strains. The YF virus Asibi strain was subcultured in embryonic mouse tissue and minced whole chicken embryo with or without nervous tissue. These passages yielded the parent 17D strain at passage level 180, 17DD at passage 195, and the 17D-204 at passage 204. 17DD was further subcultured until passage 243 and underwent 43 additional passages in embryonated chicken eggs until the vaccine batch used for 17DD virus purification (passage 284). The 17D-204 was further subcultured to produce Colombia 88 strain which, upon passage in embryonated chicken eggs, gave rise to vaccine seed lots currently in use in France (I. Pasteur, at passage 235) and in the USA (Connaught, at passage 234). Each of these 17D-204 strains was plaque purified in different cell lines, and plaque-purified virus was amplified in SW13 cells and used for cDNA cloning and sequence analyses. These 17D-204 substrains are named F-204 and C-204, respectively. The 17D-213 strain was derived from 17D-204 when the primary seed lot (S1 112-69) from the Federal Republic of Germany (FRG 83-66) was used by the World Health Organization (WHO) to produce an avian leukosis virus-free 17D seed (S1 213/77) at passage 237. This 213/77 seed was used to prepare a primary seed at the Oswaldo Cruz Foundation (FIOCRUZ S1) which was passed once more in cultured chicken embryo fibroblasts to produce experimental vaccine batches. The 17D-213 at passage 239 was tested for monkey neurovirulence and was the subject of sequence analyses together with 17DD (at passage 284) and comparison to previously published nucleotide sequences of Asibi (4) and 17D-204 (C-204[11]; F-204[14])

The YF sequences were retrieved from EMBL GenBank using the following accession numbers: 17DD U17066; 17D-213 U17067; 17D-204 France X15062; 17D-204 (ATCC) was as published[11].

Table 1 shows the total of nucleotide differences found along the genome of YF 17D virus and 17DD and 17D-213 substrains as compared to the parental virulent Asibi virus strain sequence. A total of 85 differences leading to 35 amino acid changes were detected among the three YF virus strains with the most variable areas at the amino acid level being the E protein (3.04%), followed by NS2A (2.23%) and NS4B (1.2%). On the other hand the proteins which displayed the lowest number of amino acid changes were C, NS5, NS3 and NS1 in ascending order.

Table 2 displays nucleotide and amino acid sequence comparison among all five YF virus strains and substrains. A total of 91 nucleotide differences were noted 48 of which were found to be present in all 17D vaccine strains, 21 were 17DD-specific, 13 were common to all 17D-204 (204C, 204F and 213-specific, or AB-237-specific, see *Figure 1*), four were 17D-204C-specific, one was 17D-204F-specific, three were 17D-213-specific and one was common to each 17D-204C/F and 17D-204F/213 and 17DD. Therefore, 43 nucleotide sequence differences were observed among the four YF 17D substrains a value which is close to the 48 nucleotide alterations as compared to the Asibi genome. The rate of nucleotide/amino acid changes accumulated per passage was similar considering the overall values for 17D virus (down to passage 195) and the variability among 17D virus substrains (*Table 3*) suggesting the continued evolution of these viruses.

Table 2 and *Table 3* show that from the 21 nucleotide substitutions unique to the DD genome sequence only five led to amino acid substitutions whereas six out of 13 nucleotide substitutions unique to the three 17D-204 strains sequenced did change the amino acid. Considering all changes observed for the 204 substrains (C, F and 213), and that includes each number of individual passages, 22 nucleotide changes and nine

Table 1 Total of nucleotide and amino acid changes in the genome of YF vaccine virus strains as compared to Asibi virus sequence

Virus		17D[a]		17DD		17D-213		Total	
Region	number (nt/aa)	changes (nt/aa)	% Changes (nt/aa)	Changes (nt/aa)	% Changes (nt/aa)	Changes (nt/aa)	% Changes (nt/aa)	Changes (nt/aa)	% Changes (nt/aa)
5′	118/–	0/–	0/–	0/–	0/–	0/–	0/–	0/–	0/–
C	364/121	1/–	0.27/–	0/–	0/–	1/–	0.3/–	2/–	0.6/–
prM/M	492/164	1/1	0.2/0.6	1/–	0.2/–	1/–	0.2/–	3/1	0.6/0.6
E	1479/493	11/8	0.74/1.6	5/3	0.33/0.6	4/4	0.27/0.8	20/15	1.34/3.04
NS1	1059/353	4/2	0.38/0.56	0/–	0/–	0/–	0/–	4/2	0.38/0.56
NS2A	672/224	5/4	0.74/1.78	2/–	0.3/–	2/1	0.3/0.45	9/5	1.34/2.23
NS2B	390/130	3/1	0.8/0.8	1/–	0.26/–	0/–	0/–	4/1	1.03/0.8
NS3	1869/623	5/1	0.27/0.16	7/1	0.37/0.16	4/1	0.21/0.16	16/3	0.9/0.5
NS4A	447/149	3/1	0.67/0.67	1/–	0.22/–	0/–	0/–	4/1	0.89/0.67
NS4B	750/250	2/2	0.26/0.8	0/–	0/–	2/1	0.26/0.4	4/3	0.53/1.2
NS5	2607/869	9/1	0.35/0.12	4/1	0.15/0.12	1/1	0.04/0.12	14/4	0.53/0.35
3′	621/–	4/–	0.64/–	0/–	0/–	1/–	0.16/–	5/–	0.80/–
Total	10862/3411	48/22	0.44/0.64	21/5	0.19/0.15	16/8	0.14/0.23	85/35	0.78/1.03

[a]The 17D designation represents mutations common to all 17D substrains analysed[10]

amino acid alterations were observed leading to ratios of accumulated changes of 0.35/0.16 for 17D-204 (AB 237) and 0.36/0.15 for all 17D-204 substrains. When compared to 17DD (0.22/0.06) these represent values 40% higher for nucleotide substitutions and at least 2.6 fold higher rate of amino acid substitutions suggesting a higher degree of amino acid substitution for the 17D-204 lineage (*Table 3*).

It can be argued that the 17D-204 F and C sequences are not comparable since they may contain

Table 2 Nucleotide sequence differences among YF virus 17D vaccine substrains and comparison to the corresponding positions in the parental virulent Asibi virus genome

Substrain specificity	Nucleotide/gene	Asibi	204C	204F	213	DD	Amino acid substitution
17DD	643/prM	A	A	A	A	G	–
	1436/E	G	G	G	G	A	G = D;A = S
	1437/E	A	A	A	A	G	A = G; G = S
	1558/E	C	C	C	C	A	–
	2110/E	G	G	G	G	A	–
	2220/E	C	C	C	C	T	C = T; T = V
	3599/NS2A	T	T	T	T	C	–
	3637/NS2A	C	C	C	C	T	–
	4204/NS2B	C	C	C	C	T	–
	4942/NS3	A	A	A	A	G	–
	4957/NS3	C	C	C	C	T	–
	4972/NS3	G	G	G	G	A	–
	5115/NS3	A	A	A	A	G	A = Q; G = R
	5225/NS3	A	A	A	A	C	–
	5362/NS3	C	C	C	T	A	–
	6070/NS3	C	C	C	C	T	–
	6514/NS4A	T	T	T	T	C	
	7975/NS5	C	C	C	C	T	
	8029/NS5	T	T	T	T	C	
	8808/NS5	A	A	A	A	G	A = N;G = S
	9397/NS5	A	A	A	A	G	
17D-204C	370/C	T	C	C	C	T	–
17D-204F	883/M	A	G	G	G	A	–
17D-213	1140/E	C	T	T	T	C	C = A;T = V
	1946/E	C	T	T	T	C	C = P;T = S
	2219/E	G	A	A	A	G	G = A; A = T
	4013/NS2A	C	T	T	T	C	C = L; T = F
	4054/NS2A	C	T	T	T	C	–
	4612/NS3	T	C	C	C	T	–
	4873/NS3	T	G	G	G	T	–
	5153/NS3	A	G	G	G	A	A = I; G = V
	7571/NS4B	C	A	A	A	C	
	7701/NS5	A	G	G	G	A	A = Q; G = R
	10550	T	C	C	C	T	
17D-204C	6529/NS4A	T	C	T	T	T	
	6758/NS4A	A	G	A	A	A	A = I; G = V
	9605/NS5	A	G	A	A	A	A = N; G = D
	10454	A	G	A	A	A	
17D-213	1431/E	A	A	A	C	A	A = N;C = T
	5362/NS3	C	C	C	T	A	–
	7496/NS4B	T	T	T	C	T	
17D-204F	10722	G	G	A	G	G	
204F/213/DD	7319/NS4B	G	G	A	A	A	G = E; A = K
17D-204C/F	5641/NS3	G	A	A	G	G	–

Table 3 Genetic variability among YF vaccine virus strains

	17D	All substrains[b]	DD	17D-204 (AB237)	17D-204C, F and 213	17D-213
Total changes (nt/aa)	48/22	43/15	21/5	13/6	22/9	16/8
Number of passages[a]	195	149	89	37	60	44
Average fixed changes/passage (nt/aa)	0.25/0.11	0.28/0.10	0.24/0.06	0.35/0.16	0.36/0.15	0.36/0.18

[a]Based on passage history depicted in *Figure 1*
[b]All substrains means that all of the changes found for each substrain (DD, 213, 204C, 204F, *Table 2*) was added. The number of passages corresponds to the sum of passages for each substrain

errors due to sequencing of individual clones. Moreover, the 17DD and 17D-213 viruses have no intervening passages in mammalian cells, whereas the other viruses were passaged in Vero or SW13 cells, which may have led to bonus mutations. Accordingly, if the comparison is focused on 17DD and 17D-213, for which comparability of passage history and sequencing methodology is high, a more accurate estimate of the accumulation of mutations in the viral genome could be highlighted. *Table 3* shows that again the 17D-213 virus alone has a higher frequency of mutation fixation than does YF 17DD virus.

DISCUSSION

Several YF 17D strain viruses have been used for YF vaccine production between 1937 and 1942[5] and since the establishment of the seed lot system at that time two main 17D substrains have been used for vaccine production in different countries and therefore have a different passsage history, representing two totally different branches of YF 17D virus passaging, the 17DD and 17D-204 substrains. The nucleotide sequence determination of the genome of YF virus vaccine strains 17DD and 17D-213 have allowed us to narrow down the changes previously postulated as related to viral attenuation[10] and also revealed the extent of genetic variation among vaccine strains. This issue of genetic variability of YF 17D virus had been approached earlier with the use of RNA fingerprinting technology[9]. It was shown that all 17DD differed from the 17D-204 by the absence of one RNase-generated oligonucleotide and that alteration was assumed to have occurred during the additional 40 passages of 17DD in embryonated eggs[9]. In addition the 17DD virus from Senegal and Brazil differed by another oligonucleotide possibly due to one more egg passage which separates both whereas all 17D-204 substrains, with the exception of the virus used in South Africa, had the same oligonucleotide profile. That apparently led to the conclusion that the 17D-204 lineage would be genetically more stable than 17DD[9]. Considering that such technique allows the analysis of a very limited part of the total genome, the comparison of the complete nucleotide sequences among several YF 17D vaccines substrains allows a better view of the true extent of genetic variability among these viruses and permits a more accurate determination of the mutation rates among these substrains. With regard to mutation rates, it is noteworthy that the average number of changes fixed per passage for YF 17DD is 60% that of 17D-204 in terms of nucleotide and one quarter in terms of amino acids suggesting that 17DD is indeed genetically more stable than the 17D-204 substrain.

The fact that 17DD virus contains relatively more mutations in some genes such as E, NS3 and NS5 as compared to 17D-213 would suggest that 17DD is more divergent from Asibi due to additional passages during its evolution (284 for 17DD as opposed to 239 for 17D-213). The rate at which mutations are fixed in the viral genome is what appears to distinguish both substrains.

Recently, the propagation of RNA virus vaccine strains as DNA with its correspondingly lower mutation frequency has been proposed as a method to improve the reliability of YF live attenuated vaccines. However, primate testing of the virus recovered from the 204 substrain cDNA[8] showed that this virus had a slightly higher pathological score than the 17DD suggesting caution regarding human use of the cDNA-derived virus. It is also noteworthy that the 204 substrain, used for the preparation of the cDNA library and the infectious cDNA[7,11] is closely related in terms of lineage and passage number to other 204 substrains which brought about clinical symptoms in vaccinees[12,13]. In Brazil, the 17DD substrain was by far, the most widely used substrain for both inocula and vaccine production, and a derivative of this, the EPlow substrain is still being used today. This substrain has been used for half a century for YF vaccine production in Brazil with excellent records of safety and efficacy and the results presented here show that the 17DD lineage exhibits a lower degree of mutation during its passages than the 204 lineage members. It would be of interest to try to derive an infectious cDNA which would be DD-like in its genomic sequence. This DD-like cDNA could become a stable repository for the genome of the YF vaccine virus and the sequence comparison presented here is the basis for the genetic modification and improvement of the present YF infectious cDNA.

ACKNOWLEDGEMENTS

This research was supported by grants from CNPq, Fundação Oswaldo Cruz (Programa de Apoio à Pesquisa-PAPES/FIOCRUZ) and the World Health Organization (UNDP/WHO Global Programme for Vaccines).

REFERENCES

1 Chambers, T.J., Hahn, C.S., Galler, R. and Rice, C.M. Flavivirus genome organization, expression and evolution. *Annual Reviews in Microbiology* 1990, **44**, 649.
2 Stokes, A., Bauer, J.H. and Hudson, N.P. The transmission of yellow fever to *Macacus rhesus*, preliminary note. *American Journal of Tropical Medicine and Hygiene* 1928, **8**, 103.
3 Theiler, M. and Smith, H.H. The effect of prolonged cultivation *in vitro* upon the pathogenicity of yellow fever virus. *Journal of Experimental Medicine* 1937, **65**, 787.

4 Hahn, C.S., Dalrymple, J.M., Strauss, J.H. and Rice, C.M. Comparison of the virulent Asibi strain of yellow fever virus with the 17D vaccine strain derived from it. *Proceedings of the National Academy of Science USA* 1987, **84,** 2029.
5 Fox, J.P. and Penna, H.A. Behavior of 17D yellow fever virus in rhesus monkeys. *American Journal of Hygiene* 1943, **38,** 52.
6 Monath, T.P., Brinker, K.R., Chandler, F.W., Kemp, G.E. and Cropp, C.B. Pathophysiologic correlations in a rhesus monkey model of yellow fever. *American Journal of Tropical Medicine and Hygiene* 1981, **30,** 431.
7 Rice, C.M., Grakoui, A., Galler, R. and Chambers, T.J. Transcription of infectious yellow fever RNA from full-length templates produced by *in vitro* ligation. *The New Biologist* 1989, **1,** 285.
8 Marchevsky, R.S., Mariano, J. and Ferreira, V.S. *et al.* Phenotypic analysis of yellow fever virus derived from complementary DNA. *American Journal of Tropical Medicine and Hygiene* 1995, **52,** 75.
9 Monath, T.P., Kinney, R., Schlesinger, J.J., Brandriss, M.W. and Bres, P. Ontogeny of yellow fever 17D vaccine: RNA oligonucleotide fingerprint and monoclonal antibody analysis of vaccines produced world-wide. *Journal of General Virology* 1983, **64,** 627.
10 Duarte dos Santos, C.N.D., Post, P.R., Carvalho, R., Ferreira, I.I., Rice, C.M. and Galler, R. Complete nucleotide sequence of yellow fever virus vaccine strains 17DD and 17D-213. *Virus Research* 1995, **35,** 35.
11 Rice, C.M., Lenches, E., Eddy, S.R., Shin, S.J., Sheets, R.L. and Strauss, J.H. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. *Science* 1985, **229,** 726.
12 Schoub, B.D., Dommann, C.J., Johnson, S., Downie, C. and Patel, P.L. Encephalitis in a 13-year old boy following 17D YF vaccine. *Journal of Infection* 1990, **21,** 105.
13 Merlo, C., Steffen, R., Landis, T., Tsai, T. and Karabatsos, N. Possible association of encephalitis and a 17D YF vaccination in a 29-year old traveller. *Vaccine* 1993, **11,** 691.
14 Despres, P., Cahour, A. and Dupuy, A. *et al.* High genetic stability of the coding region for the structural proteins of yellow fever virus strain 17D. *Journal of General Virology* 1987, **68,** 2245.