ELSEVIER

Virus Research 35 (1995) 35–41

Virus Research

Complete nucleotide sequence of yellow fever virus vaccine strains 17DD and 17D-213 [1]

Claudia N. Duarte dos Santos [a], Paulo. R. Post [a], Ricardo Carvalho [b], Idevaldo I. Ferreira [a], Charles M. Rice [c], Ricardo Galler [a,*]

[a] *Fundação Oswaldo Cruz, Departamento de Bioquimica e Biologia Molecular, Rio de Janeiro, RJ, Brazil 21045-900*

[b] *Fundação Oswaldo Cruz, Instituto de Tecnologia em Imunobiológicos, Rio de Janeiro, RJ, Brazil 21045-900*

[c] *Washington University School of Medicine, Department of Molecular Microbiology, 660 S. Euclid Ave., Box 8230, Saint Louis, MO 63110, USA*

Received 14 August 1994; accepted 2 September 1994

Abstract

The complete nucleotide sequence of the genome from two yellow fever (YF) virus vaccine strains, 17DD and 17D-213, has been 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 common to all 17D sub-strains. This is a significant reduction from the 67 nucleotide changes originally reported as being 17D-specific and potentially related to viral attenuation. The 48 changes are scattered throughout the genome, 26 of which are silent and 22 led to amino acid substitutions. These 22 changes are *bona fide* candidates to test by mutating the infectious YF cDNA to investigate their role in viral attenuation.

Keywords: Yellow fever virus; Flavivirus; Viral attenuation

The *Flaviviridae* is a family composed of about 70 viruses (Calisher et al., 1989). The yellow fever virus (YF) 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

* Corresponding author.

[1] Accession numbers: 17DD: GenBank U17066; 17D-213: GenBank U17067.

SSDI 0168-1702(94)00076-X

have led to new insights of genome structure and replication (reviewed by Chambers et al., 1990). The isolation of YF virus (Stokes et al., 1928) allowed the later development of an attenuated virus (17D, Theiler and Smith, 1937) 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the genomic sequences of the Asibi strain with that of the 17D-204 virus (Hahn et al., 1987; A. Grakoui and C.M. Rice, unpublished) revealed 67 and 31 nucleotide/amino acid sequence changes, respectively, scattered along the genome. The availability of an animal system that reflects human infection (Fox and Penna, 1943; Monath et al., 1981) and YF infectious cDNA (Rice et al., 1989) with an established attenuated phenotype (Marchevsky, R.S. et al., in preparation) should allow the identification of genetic determinants of viral virulence and attenuation. The approach to determine the changes responsible for attenuation would be 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. 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.

The passage history of YF virus 17DD and 213 strains has been presented elsewhere (Post et al., 1992). Briefly, The YF 17D strain was isolated after 180 passages of the virulent parental Asibi virus in tissue culture. At passage 195 the 17DD was isolated and independently passaged another 48 times in tissue culture when it was propagated in embryonated eggs until its current passage, 286. YF 17DD virus recovered from embryo homogenate has been administered to humans for half a century with excellent record of safety and immunogenicity. The 17D-213 strain, currently at passage 240, is a derivative of 17D-204 which was also derived from the virulent Asibi virus at passage 204 and thereafter independently passaged. The two 17D-strains have been used worldwide for the production of the vast majority of YF vaccine. The 17D-213 strain is a leukosis-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 virus used in this study corresponded to the medium of primary cultures of chicken embryo fibroblasts which after inoculation of monkeys was shown to have an attenuated phenotype. Aiming at the definition of genetic determinants of YF virus aftenuation, 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 and compared to other sequences of YF virus including the parental virulent Asibi virus. 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 here for 17DD and 17D-213 should be representative of the viral populations for which the attenuated phenotype has been established. We also know that using this approach PCR-related changes are not detectable since sequencing the complete E gene on both strands of 17D-204 using this methodology resulted in the same previously published nucleotide sequence (Rice et al.,

1985; Post et al., 1992). Thus, we believe that the differences noted do reflect genetic variability among the studied YF vaccine virus strains.

Virus purification and RNA extraction were performed as described (Post et al., 1991; Rico-Hesse et al., 1987). The methodology for cDNA synthesis, PCR amplification and sequencing with specific primers has been described elsewhere (Post et al., 1992). A total of 81 YF sequence-specific primers were used for PCR amplification of cDNA and complete sequence determination directly on PCR products. In all areas in which compressions were noted the other strand was sequenced. In addition, pGEM2 (Promega) cloning of PCR products was done to allow sequencing in both strands and across compressed areas. A total of 37 kb of sequence data was derived for YF 17DD of which 18.5 kb originated from direct sequencing of PCR products and 19.5 kb from sequencing plasmid DNA clones. A total of 61% of the 17DD genome was sequenced on both strands but areas for which the sequence was determined on one strand were easily readable and represents a population sequence derived from direct sequencing of reverse transcriptase-PCR products. For YF 17D-213 a total of 27.8 kb of sequence data was derived with 20.3 kb from PCR products and 7.5 kb from plasmid DNA clones leading to 60% of the genome sequenced on both strands.

Table 1 lists the 48 nucleotide differences that were found in all 17D vaccine strains when compared to the parental virulent Asibi virus sequence and therefore, possibly related to viral attenuation. These changes are scattered throughout the genome. Higher variability might be expected in the noncoding regions of the viral genome (Hahn et al., 1987; Smith and Inglis, 1987). Nevertheless, no changes were observed in the 5′ end noncoding region of any vaccine virus as compared to the virulent parental Asibi virus strain. The 4 nucleotide changes observed in the 3′ end noncoding region do not appear to affect any of the predicted secondary structures known to form in this area (Chambers et al., 1990) nor do they fall in repeated or conserved sequences shared among mosquito-borne flaviviruses.

From the 48 nucleotide sequence changes shown to be 17D-specific, 26 are silent mutations and 22 led to amino acid substitutions (Table 1). No changes were observed in the C protein gene whereas only one amino acid substitution was noted in the prM/M gene at amino acid 36 (from the amino terminus of the protein and at nucleotide 854 from the RNA 5′ end). Eleven nucleotide substitutions were observed in the E protein gene leading to 8 amino acid changes at positions 52, 170, 173, 200, 299, 305, 331 and 380 (nucleotides 1127, 1482, 1491, 1572, 1870, 1887, 1965 and 2112, respectively). Alterations at amino acids 52 and 200 are located in domain A of E protein which is conserved among flaviviruses and contains crossreactive epitopes (Mandl et al., 1989). Alterations at amino acids 299, 305, 331 and 380 are located in the B domain of E, which may be involved in the viral interaction with cell receptors as suggested by studies of epitope mapping with monoclonal antibodies (Mandl et al., 1989). Holzmann et al. (1990) using escape mutants have mapped a virulence determinant of tick-borne encephalitis virus for mice at amino acid 386 (TBE virus sequence). In addition, Lobigs et al. (1990) identified a mutated RGD sequence motif (at amino acid 390) which led to the loss of virulence of Murray Valley encephalitis virus for mice. It is noteworthy

Table 1
Nucleotide and amino acid sequence changes among YF virus Asibi and 17D strains

Gene	Position (nt)	Asibi	17D	Amino acid change
C	304	G	A	
M	854	C	U	L ⇒ F
E	1127	G	A	G ⇒ R
E	1482	C	U	A ⇒ V
E	1491	C	U	T ⇒ I
E	1572	A	C	K ⇒ T
E	1750	C	U	
E	1819	C	U	
E	1870	G	A	M ⇒ I
E	1887	C	U	S ⇒ F
E	1965	A	G	K ⇒ R
E	2112	C	G	T ⇒ R
E	2356	C	U	
NSI	2687	C	U	L ⇒ F
NSI	2704	A	G	
NSI	3274	G	A	
NSI	3371	A	G	I ⇒ V
NS2A	3613	G	A	
NS2A	3860	A	G	M ⇒ V
NS2A	4007	A	G	T ⇒ A
NS2A	4022	A	G	T ⇒ A
NS2A	4056	C	U	S ⇒ F
NS2B	4387	A	G	
NS2B	4505	A	C	I ⇒ L
NS2B	4507	U	C	
rNS3	5194	U	C	
NS3	5431	C	U	
NS3	5473	C	U	
NS3	6013	C	U	
NS3	6023	G	A	D ⇒ N
NS4A	6448	G	U	
NS4A	6829	U	C	
NS4A	6876	U	C	V ⇒ A
NS4B	7171	A	G	I ⇒ M
NS4B	7580	U	C	Y ⇒ H
NS5	7642	U	C	
NS5	7945	C	U	
NS5	8008	U	C	
NS5	8629	C	U	
NS5	10142	G	A	E ⇒ K
NS5	10243	G	A	
NS5	10285	U	C	
NS5	10312	A	G	
NS5	10338	C	U	P ⇒ L
3′	10367	U	C	
3′	10418	U	C	
3′	10800	G	A	
3′	10847	A	C	

that an RGD motif was found for all 3 YF vaccine virus strains (at amino acid 380). This sequence motif is known to mediate a number of cell interactions including receptor binding (Ruoshlati and Pierschbacher, 1987). Alteration at amino acids 170 and 173 in domain C maps very close to the position that a neutralization epitope was identified for TBE virus (Mandl et al., 1989).

The nucleotides at positions 52, 173, 200, 305 and 380 observed in the Asibi strain sequence are identical to two other wild type YF isolates, the French Viscerotropic virus (FVV, Jennings et al., 1993) and 1899/81 virus (Ballinger-Crabtree and Miller, 1990). Asibi, 1899/81 and FVV have different passage histories since their isolation but FVV is believed to have originated from the same epidemic in West Africa in 1927 as the Asibi strain whereas the 1899/81 strain was isolated from a human case of YF in Peru (South America) in 1981. Thus, this wide geographic and temporal separation of Asibi/FVV and 1899/81 lends further support to the suggestion that the 5 changes among Asibi and the 17D strains observed in the positions above observed may be relevant to attenuation of the latter.

A single amino acid change was identified in the amino terminus of NS1 (L → F). Both Asibi and 1899/81 strains have the same amino acid suggesting that this alteration could be related to attenuation of 17D virus. There are 5 nucleotide substitutions in the NS2A gene leading to amino acid changes at positions 118, 167, 172 and 183 (nucleotides 3860, 4007, 4022 and 4056). These amino acid changes did not alter the hydrophobicity profile characteristic of all NS2A proteins of flaviviruses, a characteristic which is probably related to its function. Three nucleotide substitutions were noted in the NS2B gene but only one led to an amino acid substitution at position 109 (nucleotide 4505) which again did not alter the hydrophobicity profile characteristic of all NS2B proteins of flaviviruses.

There were 5 nucleotide substitutions in the NS3 gene but only one led to an amino acid substitution at position 485 (nucleotide 6023). This protein has been shown to be multifunctional. The amino terminal domain of NS3 is a serine-protease involved in multiple cleavages in the viral polyprotein (Chambers et al., 1990). The C-terminal two-thirds of NS3 encodes a helicase activity which is believed to participate in the replication of the viral RNA and a RNA triphosphatase activity which may be involved in capping of the viral RNA (Wengler and Wengler, 1993). The change at position 485 (D → N) was found in an area which is likely to link the domains of the helicase and RNA triphosphatase as proposed in (Wengler and Wengler, 1993). In all flaviviruses for which nucleotide sequences are available, with the exception of dengue virus types 2, 3 and 4 (which have N, K and N, respectively), the amino acid in this position is aspartic acid (D) as found in the Asibi genome.

Three nucleotide substitutions in NS4A and 2 in NS4B were noted; one led to a change in NS4A (at amino acid 146, nucleotide 6876) and two in NS4B (nucleotides 7319 and 7580, amino acids 145 and 232, respectively). None of these changes altered the hydrophobic character of this protein, as discussed for NS2A/2B. Nine nucleotide substitutions were noted in the NS5 gene with two (nucleotides 10 142

and 10338) leading to amino acid changes at positions 836 (E → K) and 900 (P → L) from the amino terminus of NS5.

It is concluded that from the 31 amino acid changes found to be 17D-specific, as originally proposed (Hahn et al., 1987), only 22 have been found to be common to the 3 YF 17D virus strains (17D-204, 17D-213 and 17DD). Although all 48 nucleotide changes, both silent and coding, may also be important in attenuation, fourteen of the 17D-specific changes are arguably nonconservative regarding the character of the substituting amino acid and may be biologically more significant. However, direct testing using recombinant DNA technology to construct progeny viruses bearing specific mutations and their phenotypic characterization in appropriate animal systems will be necessary to fully assess the role of each of these mutations in viral attenuation.

Acknowledgements

We thank the Yellow Fever Vaccine Production Unit of FIOCRUZ for their constant support and for generously providing the YF vaccine preparations. This investigation received financial support from the WHO/UNDP Global Programme for Vaccines. R.G. and I.I.F. were also supported by CNPq.

References

Ballinger-Crabtree, M.E. and Miller, B.R. (1990) Partial nucleotide sequence of south america yelow fever virus strain 1899/81: structural proteins and NS1. J. Gen. Virol. 71, 2115–2121.

Calisher, C.H., Karabatsos, N., Dalrymple, J.M., et al. (1989) Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal sera. J. Gen. Virol. 70, 37–43.

Chambers, T.J., Hahn, C.S., Galler, R. and Rice, C.M. (1990) Flavivirus genome organization, expression and evolution. Annu. Rev. Microbiol. 44, 649–688.

Fox, J.P. and Penna, H.A. (1943) Behavior of 17D yellow fever virus in Rhesus monkeys. Relation to substrain, dose and neural or extraneural inoculation. Am. J. Hyg. 38, 52–172.

Hahn, C.S., Dalrymple, J.M., Strauss, J.H. and Rice, C.M. (1987) Comparison of the virulent Asibi strain of yellow fever virus with the 17D vaccine strain derived from it. Proc. Natl. Acad. Sci. USA 84, 2029–2033.

Holzmann, H., Heinz, F.X., Mandl, C., et al. (1990) A single amino acid substitution in envelope E protein of tick-borne encephalitis virus leads to attenuation in the mouse model. J. Virol. 64, 5156–5159.

Jennings, A.D., Whitby, J.E., Minor, P.D. and Barrett, A.D.T. (1993) Comparison of the nucleotide and deduced amino acid sequences of the envelope protein genes of the wild type French viscerotropic strain of yellow fever virus and the live vaccine strain, French neurotropic vaccine, derived from it. Virology 192, 692–695.

Lobigs, M., Usha, R., Nesterowicz, A., et al. (1990) Host cell selection of Murray Valley encephalitis virus variants altered at an RGD sequence in the envelope protein and in mouse neurovirulence. Virology 176, 587–595.

Mandl, C., Guirakhoo, F., Holzmann, H., et al. (1989) Antigenic structure of the flavivirus envelope protein E at the molecular level using tick-borne encephalitis virus as a model. J. Virol. 63, 564–571.

Monath, T.P., Brinker, K.R., Chandler, F.W., et al. (1981) Pathophysiologic correlations in a rhesus monkey model of yellow fever. Am. J. Trop. Med. Hyg. 30, 431–443.

Post, P.R., Santos, C.N.D., Carvalho, R., et al. (1992) Heterogeneity in envelope protein sequence and N-linked glycosylation among Yellow Fever virus vaccine strains. Virology 188, 160–167.

Post, P.R., Santos, C.N.D., Carvalho, R., et al. (1991) Molecular analysis of yellow fever virus 17DD vaccine strain. Mem. Inst. Oswaldo Cruz. 86, 239–246.

Rice, C.M., Grakoui, A., Galler, R. and Chambers, T.J. (1989) Transcription of infectious yellow fever RNA from full-length templates produced by in vitro ligation. New Biol. 1, 285–296.

Rice, C.M., Lenches, E., Eddy, S.R., et al. (1985) Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. Science 229, 726–733.

Rico-Hesse, R., Pallansch, M.A., Nottay, B. and Kew, O. (1987) Geographic distribution of wild poliovirus type 1 genotypes. Virology 160, 311–322.

Ruoshlati, E. and Pierschbacher, M.D. (1987) New perspectives in cell adhesion: RGD and integrins. Science 238, 491–497.

Smith, D.B. and Inglis, S.C. (1987) The mutation rate and variability of eukaryotic viruses. an analytical review. J. Gen. Virol. 68, 2729–2740.

Stokes, A., Bauer, J.H. and Hudson, N.P. (1928) Experimental transmission of yellow fever to laboratory animals. Am. J. Trop. Med. Hyg. 8, 103–164.

Theiler, M. and Smith, H.H. (1937a) The effect of prolonged cultivation in vitro upon the pathogenicity of yellow fever virus. J. Exp. Med. 65, 767–786.

Wengler, G. and Wengler, G. (1993) The NS3 nonstructural protein of flavivirus contains an RNA triphosphatase activity. Virology 197, 265–273.

WHO (1976) Requirements for yellow fever vaccine. WHO Expert Committee on Biological Standardization, 27th Report, Technical Series 594. WHO, Geneva, pp. 23–50.