Arch Virol (2001) 146: 2021–2027

Archives of Virology

Genome structure of Ebola virus subtype Reston: differences among Ebola subtypes

Brief Report

T. Ikegami[1,2], **A. B. Calaor**[3], **M. E. Miranda**[3], **M. Niikura**[1], **M. Saijo**[1], **I. Kurane**[1], **Y. Yoshikawa**[2], and **S. Morikawa**[1]

[1]Special Pathogens Laboratory, Department of Virology 1, National Institute of Infectious Diseases, Tokyo, Japan
[2]Department of Biomedical Science, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan
[3]Veterinary Research Department, Research Institute for Tropical Medicine, Muntinlupa, The Philippines

Accepted June 22, 2001

Summary. We determined the complete genome sequence of Ebola virus subtype Reston (EBO-R) in the Philippines in 1996. The deduced transcriptional signals were highly conserved among Ebola viruses except for the stop signal of L genes. The intergenic regions were composed of 4 to 7 nucleotides, and of 2 characteristic overlaps and a long intergenic region. The glycoprotein (GP) had several amino acid differences from EBO-R isolated in 1989 and 1992. The variety of GP sequences strongly suggests the independent introduction of EBO-R from unknown natural reservoirs in 1996.

*

Ebola virus (EBOV) and Marburg virus (MBGV) are nonsegmented negative strand viruses, and a member of the family *Filoviridae* in the Order *Mononegavirale*. The genome of filoviruses is composed of seven genes; nucleoprotein (NP), P protein (VP35), matrix protein (VP40), glycoprotein (GP), second nucleoprotein (VP30), protein associated with the envelope (VP24) and RNA-dependent RNA polymerase (L) [3–5, 12]. In addition, the GP of EBOV is found in two forms sharing identical NH_2 terminal sequences, i.e. secreted GP (sGP) and transmembrane GP (GP) [15, 16]. EBOVs are subdivided into four subtypes, Zaire (EBO-Z), Sudan (EBO-S), Reston (EBO-R) and Ivory Coast (EBO-IC) on the basis of the significant differences in the antigenicity and the nucleotide sequences [3, 15]. GP of each subtype differs from the other subtypes by 37–41% and 34–43% at the

nucleotide and amino acid levels, respectively [3, 15]. Because of the similarity in the amino acid sequences, the subtypes of EBOV share common epitopes [10].

Ebola virus subtype Reston (EBO-R) first emerged among monkeys in Reston, Virginia and in Pennsylvania in 1989 [7]. Epizootics in monkeys followed in Alice, Texas in 1990, in Siena, Italy in 1992 [18], and again in Alice Texas [11] and in the Philippines [8] in 1996. In all these epizootics, the monkeys were imported from the Philippines [6, 8]. During these outbreaks, it was shown that EBO-R was not pathogenic to human [2, 8]. Phylogenetic analysis of the GP region indicates that EBO-R may be indigenous in Asia and not introduced to this region from Africa in the recent past [3].

The complete nucleotide sequence was determined only for EBO-Z. In the present study, we determined the complete nucleotide sequence of EBO-R from the Philippines in 1996 (GenBank accession no. AB050936). In order to clarify the structural differences within EBOV subtypes, we compared the deduced transcriptional signals, intergenic regions, and open reading frames (ORFs) among EBOVs and MBGV. Furthermore, we investigated the nucleotide and amino acid replacements in the GP among EBO-Rs that emerged in 1989, 1992 and 1996, and demonstrated the lineage among these outbreaks.

Total RNAs were extracted from the liver of a monkey naturally infected with EBO-R (Philippines, 1996) by RNAzol B (TEL-TEST, Inc. Friendswood, TX). Random hexamer-primed, first strand cDNAs were synthesized using Ready-To-Go You-Prime First-Strand Beads (Amersham Pharmacia Biotech Inc.). We prepared the primers according to the sequences of EBO-S or EBO-Z, and then performed PCR with the first strand cDNAs using a commercial PCR kit (Expand High Fidelity PCR System, Roche Diagnostics, Mannheim, Germany). For the direct sequence of the 5′ end of viral (negative) and anti-viral (positive) sense RNAs, we performed 5′Rapid Amplification of cDNA Ends assay (5′RACE) using a commercial kit (5′RACE System for Rapid Amplification of cDNA Ends, Version 2.0, Life Technologies, Rockville, MD). The PCR products of 5′RACE reactions were used for direct sequencing on the ABI PRISM 310 Genetic Analyzer (Perkin-Elmer). The alignment of nucleotide or deduced amino acid sequence was performed on computer using the CLUSTAL W program.

The genomic RNA of EBO-R consisted of 18,890 bases. The genome RNA length was slightly shorter than that of EBO-Z (18,957 bases; GenBank accession no. AF086833) and MBGV (19,104 bases; GenBank accession no. Z12132). The structure was 3′-leader-NP-VP35-VP40-GP-VP30-VP24-L-trailer-5′, which was analogous to that of EBO-Z and MBGV.

Since we did not map the 5′ and 3′ termini of mRNAs of each viral gene, we deduced the transcriptional start (R1) and stop signals (R2) according to that of EBO-Z (Table 1). The consensus sequences of R1 and R2 in the filoviruses are 3′-CUNCNUNUAAUU-5′ and 3′-UAAUUCUUUUU-5′, respectively [3, 9]. However, the R2 in MBGV VP40 gene, 3′-UCAUUCUUUUU-5′, and that in EBO-Z L gene, 3′-UAAUAUUUUUU, are slightly different from the consensus sequences [17]. EBO-R also shares the consensus sequences of R1 and R2 in each gene. The R2 in L gene was not identical to the consensus R2 sequence, but

Table 1. Comparison of conserved transcriptional signals and intergenic regions of EBOV and MBGV

		EBO-R, 1996 (Philippines)	EBO-Z, 1976 (Zaire)	MBGV, 1980 (Kenya)
NP	Start signals	CUCCUUCUAAUU	CUCCUUCUAAUU	CUUCUUAUAAUU
	Stop signals	UAAUUCUUUUU	UAAUUCUUUUUU	UAAUUCUUUUU
	Intergenic	GUAAA	GAUUA	GAUUUUUAG
VP35	Start signals	CUACUUCUAAUU	CUACUUCUAAUU	CUUCUUAUAAUU
	Stop signals	UAAUUCUUUUU	UAAUUCUUUUU	UAAUUCUUUUU
	Intergenic	*	*	GAUA
VP40	Start signals	CUACUUCUAAUU	CUACUUCUAAUU	CUUCUUGUAAUU
	Stop signals	UAAUUCUUUUU	UAAUUCUUUUUU	UCAUUCUUUUU
	Intergenic	CUAUAUG	AGCCG	GAUUA
GP	Start signals	CUACUUCUAAUU	CUACUUCUAAUU	CUUCUUGUAAUU
	Stop signals	UAAUUCUUUUU	UAAUUCUUUUU	UAAUUCUUUUU
	Intergenic	GAAUA	*	CUA-$(N)_{91}$-CAG
VP30	Start signals	CUGCUUCUAAUU	CUACUUCUAAUU	CUUCUUGUAAUU
	Stop signals	UAAUUCUUUUUU	UAAUUCUUUUU	UAAUUCUUUUU
	Intergenic	AGU-$(N)_{124}$-GUC	GAC-$(N)_{137}$-UUA	*
VP24	Start signals	CUACUUCUAAUU	CUACUUCUAAUU	CUUCUUGUAAUU
	Stop signals	UAAUUCUUUUU	UAAUUCUUUUUU	UAAUUCUUUUU
	Intergenic	*	CGGA	AGUUA
L	Start signals	CUCCUUCUAAUU	CUCCUUCUAAUU	CUACCUAUAAUU
	Stop signals	UAAUAUUUUUU	UAAUAUUUUUU	UAAUUCUUUUU

Genebank accession numbers; EBO-R (AB050936), EBO-Z (AF086833), MBGV (Z12132)
*No intergenic region was detected because of the gene overlaps.

identical to that of EBO-Z (Table 1). Therefore, the sequence in the R2 of L gene, 3′-UAAUAUUUUUU-5′, is likely to be the common feature of EBOVs.

The location of highly conserved R1 and R2 revealed the margins of each gene. In EBO-R and EBO-Z, each gene was generally flanked by intergenic regions composed of 4 to 7 nucleotides, except for a long intergenic region (LI) between VP30 and VP24 gene and 2 overlaps (VP35-VP40 and VP24-L for EBO-R, and VP35-VP40 and GP-VP30 for EBO-Z), while MBGV has the LI between GP and VP30 gene and 1 overlap at VP30-VP24 (Table 1). One of these 2 overlaps of EBOV was common at VP35-VP40, the other was located at VP24-L for EBO-R and at GP-VP30 for EBO-Z. The location of overlapping genes has been used as a criterion to classify the isolates within the family *Filoviridae* [4]. The difference in the location of overlapping genes between EBO-Z and EBO-R indicates the diversity among EBOV subspecies.

The leader sequences of EBO-R, EBO-Z and MBGV were 53, 55 and 48 bases in length, respectively. The nucleotide sequence of 3′-extremity of EBO-R leader, 3′-CUGUGUGUUUU-5′ (11 bases), was identical to those of EBO-Z and MBGV. The complementary sequence could form the panhandle structure

Table 2. Homologies to EBOV subspecies in deduced amino acid sequence of each ORF (%)

		EBO-R (1989)	EBO-R (1992)	EBO-Z (1976)	EBO-S (1976)	EBO-IV (1994)
EBO-R (1996)	NP	ND	ND	67.7	66.8	ND
	VP35	ND	ND	64.7	ND	ND
	VP40	ND	ND	73.4	ND	ND
	GP	97.9	98.1	58.1	59.6	57.6
	VP30	ND	ND	68.8	ND	ND
	VP24	ND	ND	81.3	ND	ND
	L	ND	ND	73.9	73.5[a]	ND

EBO-S, 1976 (NP:AF173836, GP:U28134), EBO-S, 1979 (L:U23458), EBO-IV (U28006)

ND: Not determined because the sequences were not available

[a]Homology to EBO-S (1979)

or stem-loop structure [17]. It is suggested that the secondary structures act as important elements for genomic and antigenomic promoters [17]. Putative 5′ termini of mRNAs of EBO-R could also form stem-loop structure. As has been shown for EBO-Z and MBGV [9], these stable structures may serve as important regulatory elements.

It was reported that the length of 3′-untranslated region (3′UTR) of NP gene in EBO-Z and MBGV were 414 and 55 bases, respectively. Thus, 3′UTR of EBO-Z was extremely long compared to that of MBGV [4]. It should be pointed out that the length in 3′UTR of EBO-R NP gene was 408 bases, and that the long 3′UTR of NP gene in EBO-R supports the criterion for family *Filoviridae* proposed using EBO-Z and MBGV.

The deduced amino acid sequence of each gene of EBO-R was aligned in comparison with those of other EBOVs (Table 2). Homologies in amino acid sequences to EBO-Z were 67.7% for NP, 64.7% for VP35, 73.4% for VP40, 58.1% for GP, 68.8% for VP30, 81.3% for VP24 and 73.9% for L. Homologies to EBO-S were 66.8% for NP, 59.6% for GP and 73.5% for L. The central hydrophilic region of GP is suggested to be located outside the lipid membrane and be affected by the selection of host immune responses [1, 12]. On the other hand, NP, VP35, VP30 and L are associated with the viral ribonucleocapsid complex (RNP), which includes the conserved regions among EBOVs [3, 4, 12, 13]. Therefore, the GP showed the lowest homology in the seven viral proteins among EBOV subspecies.

In order to clarify the lineages of EBO-Rs isolated in 1989, 1992 and 1996, we compared the nucleotide and amino acid sequences of GP among Reston isolate in 1989 (GenBank accession no. U23152), Pennsylvania isolate in 1989 (GenBank accession no. AF034645), Italy isolate in 1992 (GenBank accession no. U23417), Philippine isolate in 1992 (GenBank accession no. U23416), and EBO-R from Philippines in 1996 (Table 3). One and 7 silent nucleotide changes were seen between Reston and Pennsylvania isolates in 1989, and between Italy

Table 3. The number of nucleotide (NT) and amino acid (AA) replacements in GP of EBO-R isolates

NT \ AA	Reston (1989)	Pennsylvania (1989)	Italy (1992)	Philippines (1992)	Philippines (1996)
Reston (1989)		0	10	10	14
Pennsylvania (1989)	1		10	10	14
Italy (1992)	20	21		0	13
Philippines (1992)	25	26	7		13
Philippines (1996)	25	26	25	30	

and Philippine isolates in 1992, respectively. No amino acid substitutions were, however, observed within the different isolates in the same year. On the other hands, 10, 13 and 14 amino acids substitutions were found between the isolates in 1989 and 1992, between 1992 and 1996 and between 1989 and 1996, respectively (Table 3). We also investigated the locations of amino acid substitutions among EBO-Rs. It is known that hydrophilic variable region exists between amino acid residues 190 and 478 within the GP [12]. In fact, 52.4% and 66.7% of the nucleotide and amino acid substitutions among EBO-R isolates were found within this region, respectively. Thirteen out of 15 nucleotide mutations involving amino acid replacement existed between amino acid residues 424 and 528. Nine silent mutations existed between the nucleotide positions 6255 and 6504, outside the variable region. While EBO-R isolates in 1989 has 13 deduced N-linked glycosylation sites within the variable region, EBO-Rs in 1992 and 1996 have 12 sites because of the substitution of asparagine (N to K and N to E, respectively) at amino acid residue 229 [14]. Serial passage of initially nonlethal EBO-Z in guinea pigs resulted in the selection of variants with high pathogenicity with the significant amino acid substitutions in VP24 [16]. Along with the fact that immune response was not observed in any monkeys during EBO-R outbreak [8], it seems likely that amino acid substitutions do not occur easily within an outbreak. Therefore, several amino acid differences among GPs of EBO-R in 1989, 1992 and 1996 strongly indicate independent introduction of EBO-R into the monkey facility from yet-identified reservoir of the virus.

Acknowledgements

We thank the staff of the Veterinary Research Department and the Molecular Biology Laboratory of the Research Institute for Tropical Medicine, The Philippines, for their technical assistance. This work was partially supported by a grant-in-aid from the Ministry of Health and Welfare of Japan, and the Human Science Foundation, Japan.

References

1. Bukreyev AA, Volchkov VE, Blinov VM, Dryga SA, Netesov SV (1995) The complete nucleotide sequence of the Popp (1967) strain of Marburg virus: a comparison with the Musoke (1980) strain. Arch Virol 140: 1589–1600
2. Centers for Disease Control (1990) Update: evidence of filovirus infection in an animal caretaker in a research/service facility. Morb Mortal Wkly Rep 39: 296–297
3. Feldmann H, Kiley MP (1999) Classification, structure, and replication of filoviruses. Curr Top Microbiol Immunol 235: 1–21
4. Feldmann H, Klenk H-D, Sanchez A (1993) Molecular biology and evolution of filoviruses. Arch Virol [Suppl] 7: 81–100
5. Feldmann H, Muhlberger E, Randolf A, Will C, Kiley MP, Sanchez A, Klenk H-D (1992) Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle. Virus Res 24: 1–19
6. Hayes CG, Burans JP, Ksiazek TG, Del Rosario RA, Miranda ME, Manaloto CR, Barrientos AB, Robles CG, Dayrit MM, Peters CJ (1992) Outbreak of fatal illness among captive macaques in the Philippines caused by an Ebola-related filovirus. Am J Trop Med Hyg 46: 664–671
7. Jahrling PB, Geisbert TW, Dalgard DW, Johnson ED, Ksiazek TG, Hall WC, Peters CJ (1990) Preliminary report: isolation of Ebola virus from monkeys imported to USA. Lancet 335: 502–505
8. Miranda ME, Ksiazek TG, Retuya TJ, Khan AS, Sanchez A, Fulhorst CF, Rollin PE, Calaor AB, Manalo DL, Roces MC, Dayrit MM, Peters CJ (1996) Epidemiology of Ebola (subtype Reston) virus in the Philippines, 1996. J Infect Dis [Suppl] 179: 115–119
9. Muhlberger E, Trommer S, Funke C, Volchkov V, Klenk H-D, Becker S (1996) Termini of all mRNA species of Marburg virus: sequence and secondary structure. Virology 223: 376–380
10. Richman DD, Cleveland PH, McCormick JB, Johnson KM (1983) Antigenic analysis of strains of Ebola virus: identification of two Ebola virus subtypes. J Infect Dis 147: 268–271
11. Rollin PE, Williams RJ, Bressler DS, Pearson S, Cottingham M, Pucak G, Sanchez A, Trappier SG, Peters RL, Greer PW, Zaki S, Demarcus T, Hendricks K, Kelley M, Simpson D, Geisbert TW, Jahrling PB, Peters CJ, Ksiazek TG (1999) Ebola (Subtype Reston) virus among quarantined nonhuman primates recently imported from the Philippines to the United States. J Infect Dis [Suppl] 179: 108–114
12. Sanchez A, Kiley MP, Holloway BP, Auperin DD (1993) Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. Virus Res 29: 215–240
13. Sanchez A, Kiley MP, Holloway BP, McCormick JB, Auperin DD (1989) The nucleoprotein gene of Ebola virus: cloning, sequencing, and in vitro expression. Virology 170: 81–91
14. Sanchez A, Ksiazek TG, Rollin PE, Miranda MEG, Trappier SG, Khan AS, Peters CJ, Nichol ST (1999) Detection and molecular characterization of Ebola virus causing disease in human and nonhuman primates. J Infect Dis [Suppl] 179: 164–169
15. Sanchez A, Trappier SG, Mahy BW, Peters CJ, Nichol ST (1996) The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. Proc Natl Acad Sci USA 93: 3602–3607
16. Volchkov VE, Chepurnov AA, Volchkova VA, Ternovoj VA, Klenk H-D (2000) Molecular characterization of guinea pig-adapted variants of Ebola virus. Virology 277: 147–155

17. Volchkov VE, Volchkova VA, Chepurnov AA, Blinov VM, Dolnik O, Netesov SV, Feldmann H (1999) Characterization of the L gene and 5′ trailer region of Ebola virus. J Gen Virol 80: 355–362
18. World Health Organization (1990) Viral haemorrhagic fever in imported monkeys. Wkly Epidemiol Rec 67: 142–143
19. Yang Z-Y, Delgado R, Xu L, Todd RF, Nabel EG, Sanchez A, Nabel GJ (1998) Distinct cellular interactions of secreted and transmembrane Ebola virus glycoproteins. Science 279: 1034–1037

Author's address: Dr. S. Morikawa, Special Pathogens Laboratory, Department of Virology 1, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama, Tokyo 208-0011, Japan; e-mail: morikawa@nih.go.jp

Received May 8, 2001