The Origin and Evolution of Ebola and Marburg Viruses

Yoshiyuki Suzuki† and Takashi Gojobori**

*Center for Information Biology, National Institute of Genetics, Mishima, Japan; and †Department of Microbiology, Akita University School of Medicine, Akita, Japan

Molecular evolutionary analyses for Ebola and Marburg viruses were conducted with the aim of elucidating evolutionary features of these viruses. In particular, the rate of nonsynonymous substitutions for the glycoprotein gene of Ebola virus was estimated to be, on the average, 3.6×10^{-5} per site per year. Marburg virus was also suggested to be evolving at a similar rate. Those rates were a hundred times slower than those of retroviruses and human influenza A virus, but were of the same order of magnitude as that of hepatitis B virus. When these rates were applied to the degree of sequence divergence, the divergence time between Ebola and Marburg viruses was estimated to be more than several thousand years ago. Moreover, most of the nucleotide substitutions were transitions and synonymous for Marburg virus. This suggests that purifying selection has operated on Marburg virus during evolution.

Introduction

Ebola and Marburg viruses are known to be the etiological agents of haemorrhagic fever, which has a high mortality rate (Martini and Siegert 1971; International Commission 1978; WHO/International Study Team 1978; Baron, McCormick, and Zubeir 1983; Centers for Disease Control and Prevention 1995), and thus have been classified as "biosafety level 4" agents (Richardson and Barkley 1988). These viruses have also been classified into the genus *Filovirus,* which is the sole member of the family *Filoviridae* (Kiley et al. 1982). The genome of these viruses is the nonsegmented, negative-stranded RNA (Regnery, Johnson, and Kiley 1981), and thus they have been further classified into the order *Mononegavirales* (Pringle 1991). Their genomes encode the same set of seven genes in the same order, namely nucleoprotein (NP), viral structural protein 35 (VP35), VP40, glycoprotein (GP), VP30, VP24, and RNA-dependent RNA polymerase (L), from the 3′ to 5′ end (Feldmann et al. 1992; Sanchez et al. 1993).

From the molecular evolutionary point of view, it is of importance to elucidate the origins and evolutionary modes of these viruses. In particular, to examine the rates and patterns of nucleotide substitutions for these viruses is important for understanding the mutation mechanism, and it is also useful for predicting the future evolution of these viruses. Such knowledge can lead us to the development of antiviral drugs and effective vaccines for Ebola and Marburg viruses. In this study, we estimated the rates of nucleotide substitutions for Ebola and Marburg viruses. Applying the estimated rates to the degree of sequence divergence, we further estimated the divergence time not only among Ebola virus strains but also between Ebola and Marburg viruses. The pattern of nucleotide substitutions is also discussed to clarify the evolutionary modes of these viruses.

Key words: Ebola virus, Marburg virus, substitution rate, divergence time, substitution pattern, purifying selection.

Address for correspondence and reprints: Takashi Gojobori, Director and Professor, Center for Information Biology, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka-ken 411, Japan. E-mail: tgojobor@genes.nig.ac.jp.

Mol. Biol. Evol. 14(8):800–806. 1997

Materials and Methods

Sequence Data Used

Sequence data for Ebola and Marburg viruses were collected from the international DNA data banks DDBJ/EMBL/GenBank. The sequence data used in this study are summarized in table 1. In our analyses, we assumed that no sequences of Ebola and Marburg viruses changed after isolation. If these viruses did change their genome sequences after isolation, the rates of nucleotide substitutions estimated in the present analysis would be underestimates, but they would still give us important information.

Rates of Nonsynonymous Substitutions for Ebola and Marburg Viruses

We first made alignments of homologous sequences for Ebola and Marburg viruses using CLUSTAL W (Thompson, Higgins, and Gibson 1994). Then, the neighbor-joining method (Saitou and Nei 1987) was used for constructing phylogenetic trees, with the distances estimated using the method of Nei and Gojobori (1986). The reliability of clustering in the phylogenetic trees was tested by the bootstrap method with 1000 replications (Felsenstein 1985). Phylogenetic trees were constructed for all genes for the number of nonsynonymous substitutions. Unfortunately, we could not construct phylogenetic trees for synonymous substitutions, because the number of synonymous substitutions among different virus strains of Ebola virus or between Ebola and Marburg viruses was so large that we could not estimate it accurately. We also could not estimate the substitution rate of Ebola virus for all genes except the GP gene, because sequence data for Ebola virus were available for only one strain in other genes. The rates were estimated from the phylogenetic trees by dividing the difference in branch lengths of two strains of interest from the last common ancestor by the difference in their isolation times. In the case of the GP gene of Ebola virus, however, we excluded the sequences of Marburg virus in constructing the phylogenetic tree, because we could estimate branch lengths accurately by using only Ebola virus strains. Then, we made comparisons between all possible pairs of viral sequences from the same subtype to avoid getting large variances.

Table 1
Sequence Data Used in this Study

Gene	Accession Number	Virus	Place and Time	Codon Numbers	Reference
NP.....	L11365	Ebola	Yambuku, 1976	20–409	Sanchez et al. (1989)
	M72714	Marburg	Kenya, 1980	2–391	Sanchez et al. (1992)
	Z29337	Marburg	Marburg, 1967	2–391	Bukreyev et al. (1995*b*)
VP35.....	L11365	Ebola	Yambuku, 1976	78–339	Sanchez et al. (1993)
	X61274	Ebola	Yambuku, 1976	78–340	Bukreyev et al. (1993*a*)
	Z12132	Marburg	Kenya, 1980	67–329	Feldmann et al. (1992)
	Z29337	Marburg	Marburg, 1967	67–329	Bukreyev et al. (1993*a*)
VP40.....	L11365	Ebola	Yambuku, 1976	67–295	Sanchez et al. (1993)
	X61274	Ebola	Yambuku, 1976	67–295	Bukreyev et al. (1993*a*)
	Z12132	Marburg	Kenya, 1980	55–283	Feldmann et al. (1992)
	Z29337	Marburg	Marburg, 1967	55–283	Bukreyev et al. (1993*a*)
GP.....	U23069	Ebola	Nzara, 1979	25–185, 511–672	Sanchez et al. (1996)
	U23152	Ebola	Reston, 1989	26–186, 512–673	Sanchez et al. (1996)
	U23187	Ebola	Yambuku, 1976	25–185, 511–672	Sanchez et al. (1996)
	U23416	Ebola	Manila, 1992	26–186, 512–673	Sanchez et al. (1996)
	U23417	Ebola	Siena, 1992	26–186, 512–673	Sanchez et al. (1996)
	U28006	Ebola	Tai, 1994	25–185, 511–672	Sanchez et al. (1996)
	U28077	Ebola	Kikwit, 1995	25–185, 511–672	Sanchez et al. (1996)
	U28134	Ebola	Maridi, 1976	25–185, 511–672	Sanchez et al. (1996)
	U31033	Ebola	Yambuku, 1976	25–185, 511–672	Volchkov et al. (1995)
	Z12132	Marburg	Kenya, 1980	11–169, 512–673	Feldmann et al. (1992)
	Z29337	Marburg	Marburg, 1967	11–169, 512–673	Bukreyev et al. (1993*b*)
VP30.....	L11365	Ebola	Yambuku, 1976	62–159, 167–254	Sanchez et al. (1993)
	Z12132	Marburg	Kenya, 1980	67–166, 173–260	Feldmann et al. (1992)
	Z29337	Marburg	Marburg, 1967	67–166, 174–261	Bukreyev et al. (1995*a*)
VP24.....	L11365	Ebola	Yambuku, 1976	2–251	Sanchez et al. (1993)
	Z12132	Marburg	Kenya, 1980	2–253	Feldmann et al. (1992)
	Z29337	Marburg	Marburg, 1967	2–253	Bukreyev et al. (1995*a*)
L.....	U23458	Ebola	Nzara, 1979	5–1161, 1163–1650, 1784–2209	Not published
	Z12132	Marburg	Kenya, 1980	2–1163, 1189–1674, 1903–2328	Muehlberger et al. (1992)
	Z29337	Marburg	Marburg, 1967	2–1163, 1189–1674, 1903–2328	Bukreyev et al. (1995*b*)

NOTE.—This table shows genes, accession numbers of the sequence data in DDBJ/EMBL/GenBank, virus names, places and times of outbreaks, codon numbers of gene regions examined, and references. Several gaps were conducted in the analyzed regions in sequence alignments.

The divergence times among Ebola virus strains and between Ebola and Marburg viruses were estimated on the assumption that these viruses had evolved at almost the same substitution rate.

Patterns of Nucleotide Substitutions for Marburg Virus

The pattern of nucleotide substitutions was examined for two Marburg virus strains for which the entire genome sequences were available (DDBJ/EMBL/GenBank accession numbers M72714 and Z12132, and Z29337). All nucleotide changes between them were assumed to have occurred through single-nucleotide substitutions. This assumption is reasonable because the nucleotide sequences of these strains were closely related (94%–97% identity). The numbers of substitutions between two particular nucleotides were summed up, and the values thus obtained were corrected by base compositions using the method of Gojobori, Li, and Graur (1982). The corrected values represent the substitution numbers from a particular nucleotide to another one in 100 nucleotides of a hypothetical sequence which contains equal amounts of the four nucleotides.

Results and Discussion

Rates of Nonsynonymous Substitutions for Ebola and Marburg Viruses

The rates of nonsynonymous substitutions for Ebola and Marburg viruses are summarized in tables 2 and 3, respectively. Unfortunately, we could not estimate the rate of synonymous substitutions because the number of synonymous substitutions among different virus strains of Ebola virus or between Ebola and Marburg viruses was so large that we could not estimate it accurately. We also could not estimate the substitution rate of Ebola virus for all genes except the GP gene, because sequence data for Ebola virus were available for only one strain in other genes.

For Ebola virus, the average rate of nonsynonymous substitutions for the GP gene was estimated to be 3.6×10^{-5} per site per year (table 2). The value of standard error appeared to be relatively large. This might be due to the relatively small difference in isolation times compared with the slow rate of nucleotide substitutions. However, the rate was estimated to be at the same order in almost all comparisons, as shown in table 2. Negative values were obtained in the estimations for the NP, VP40, GP, and L genes of Marburg virus, which would be due to statistical fluctuations because of the relatively large distances between Ebola and Marburg viruses. However, the values for VP35, VP30, and VP24 indicate that Marburg virus is evolving at the rate of 10^{-5} to 10^{-4} per site per year. These rates were compared with those of other RNA viruses and mammals in table 4. Most of the RNA viruses are known to evolve

Table 2
Rate of Nonsynonymous Substitutions for the GP Gene of Ebola Virus

Strains Compared[a]	Difference in Branch Lengths (× 10^{-4}/site)	Difference in Isolation Times (years)	Rate (× 10^{-4}/site/year)
U23187 and U28077 ...	7.16	19	0.38 ± 1.01
U31033 and U28077 ...	7.16	19	0.38 ± 1.01
U23152 and U23416 ...	1.16	3	0.39 ± 6.33
U23152 and U23417 ...	2.72	3	0.91 ± 6.51
U23069 and U28134 ...	0.00	3	0.00 ± 0.00
Average	—	—	0.36 ± 1.09

NOTE.—The rates were estimated from the phylogenetic tree, excluding the sequences of Marburg virus, by dividing the difference in branch lengths of two strains of interest from the last common ancestor by the difference in their isolation times.

[a] Reference sequences used were U23069, U23152, U23416, U23417, U28006, and U28134 for the comparisons between U23187 and U28077 and between U31033 and U28077; U23069, U23187, U28006, U28077, U28134 and U31033 for the comparisons between U23152 and U23416 and between U23152 and U23417; and U23187, U23152, U23416, U23417, U28006, U28077, and U31033 for the comparison between U23069 and U28134.

at the rate of 10^{-5} to 10^{-3} per site per year, and the rates for Ebola and Marburg viruses seem to be roughly of the same order of magnitude, suggesting that these viruses share the molecular mechanisms of rapid evolution with other RNA viruses. Compared with other RNA viruses, however, these viruses seem to be evolving relatively slowly. In particular, both Ebola and Marburg viruses have substitution rates approximately a hundred times slower than retroviruses and human influenza A virus. This is consistent with the previous report that suggested genetic stability in Ebola virus from the results of oligonucleotide mapping (Cox et al. 1983). Then, we speculate the following reasons for the relatively slow rates of nonsynonymous substitutions for Ebola and Marburg viruses. First, the RNA-dependent RNA polymerase of Ebola and Marburg viruses may not

Table 3
Rates of Nonsynonymous Substitutions for Marburg Virus and Divergence Times Between Ebola and Marburg Viruses

Gene	Number of Nonsynonymous Sites	Difference in Branch Lengths (× 10^{-3}/site)	Substitution Rate (× 10^{-4}/site/year)	Divergence Time (years)
NP	905.67	NG[a]	NG	—
VP35 ...	602.67	4.68	3.60 ± 2.57	1,000
VP40 ...	523.92	NG	NG	—
GP	738.70	NG	NG	—
VP30 ...	437.33	0.49	0.38 ± 4.69	9,800
VP24 ...	583.67	1.65	1.27 ± 2.29	2,800
L	4,861.56	NG	NG	—

NOTE.—The difference in isolation times was 13 years for all comparisons.
[a] NG: negative value was obtained.

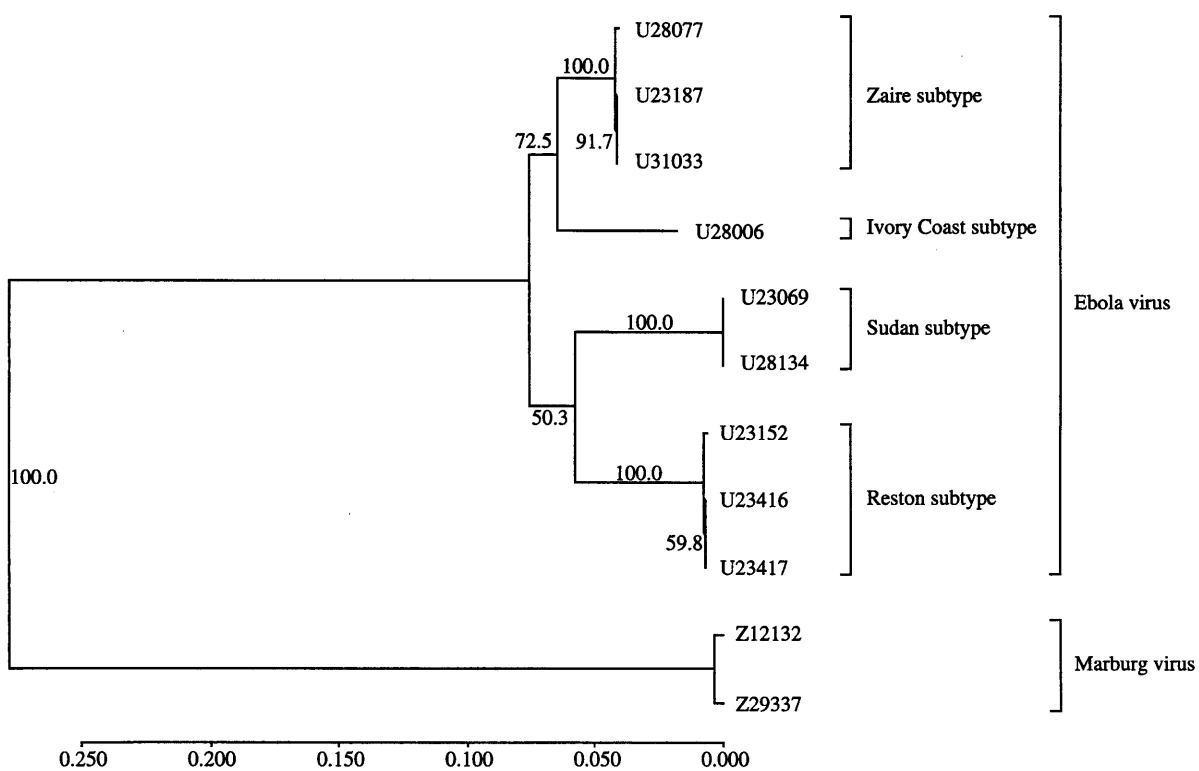

FIG. 1.—Phylogenetic tree constructed for the GP gene of Ebola and Marburg viruses by the neighbor-joining method (Saitou and Nei 1987), with distances for nonsynonymous sites estimated by the method of Nei and Gojobori (1986). The bootstrap probability for each node is also indicated (Felsenstein 1985). When we estimated the substitution rate of Ebola virus, we excluded the sequences of Marburg virus and constructed another phylogenetic tree (data not shown), in which the topology among Ebola virus strains was identical with that of the former one.

Table 4
Comparisons of the Rates of Nonsynonymous Substitutions for Ebola and Marburg Viruses with Those of Various RNA Viruses and Mammals

Virus and Organism	Gene	Substitution Rate (/site/year)	Reference
Ebola virus	GP	3.6×10^{-5}	
Marburg virus	VP35	3.6×10^{-4}	
	VP30	3.8×10^{-5}	
	VP24	1.3×10^{-4}	
HIV-1[a]	*gag*	$(1.0 - 3.9) \times 10^{-3}$	Li, Tanimura, and Sharp (1988); Gojobori, Moriyama, and Kimura (1990); Gojobori et al. (1994)
	pol	1.6×10^{-3}	Li, Tanimura, and Sharp (1988)
	env	$(3.9 - 5.1) \times 10^{-3}$	Li, Tanimura, and Sharp (1988); Gojobori et al. (1994)
	envhv	14.0×10^{-3}	Li, Tanimura, and Sharp (1988)
Human influenza A virus	HA (H3)	$(2.9 - 3.6) \times 10^{-3}$	Gojobori, Moriyama, and Kimura (1990); Hayashida et al. (1985)
	NA (N1)	3.7×10^{-3}	Hayashida et al. (1985)
	NA (N2)	2.8×10^{-3}	Hayashida et al. (1985)
MMSV[b]	v-*mos*	8.2×10^{-4}	Gojobori, Moriyama, and Kimura (1990)
MMLV[c]	*gag*	5.4×10^{-4}	Gojobori and Yokoyama (1985)
HCV[d]	C	6.3×10^{-4}	Ina et al. (1994)
	E	3.2×10^{-4}	Ina et al. (1994)
	NS1	7.5×10^{-4}	Ina et al. (1994)
	NS3	3.3×10^{-4}	Ina et al. (1994)
	NS5	2.2×10^{-4}	Ina et al. (1994)
HBV[e]	P	1.5×10^{-5}	Orito et al. (1989)
	pre-S	2.6×10^{-5}	Orito et al. (1989)
	C	1.8×10^{-5}	Orito et al. (1989)
	X	5.5×10^{-5}	Orito et al. (1989)
Mammals	α-globin	5.6×10^{-10}	Li, Luo, and Wu (1985)

[a] Human immunodeficiency virus.
[b] Moloney murine sarcoma virus.
[c] Moloney murine leukemia virus.
[d] Hepatitis C virus.
[e] Hepatitis B virus. HBV is included because it is known to replicate itself via the RNA transcript.

be so error-prone. Second, the replication frequency may be relatively low in the natural host in comparison with retroviruses and human influenza A virus. Third, the number of the reservoir of these viruses in nature may be relatively small probably because of a low level of infectivity of these viruses. Finally, strong functional constraints may be operating on these viruses during evolution, particularly on GP and VP30, for we focused only on nonsynonymous substitutions, which change the coding amino acid. When we examined the rates of synonymous substitutions for Ebola and Marburg viruses, they were estimated to be at most 1.35×10^{-2} and 1.77×10^{-2} per site per year, respectively. The rates of synonymous substitutions for retroviruses and human influenza A virus have been estimated to be 10^{-2} to 10^{-3} (Hayashida et al. 1985; Li, Tanimura, and Sharp 1988; Gojobori, Moriyama, and Kimura 1990; Gojobori et al. 1994). Thus, we could not rule out the possibility that the relatively slow rates of nonsynonymous substitutions for Ebola and Marburg viruses were due to the strong functional constraint while the mutation rates were as high as those of retroviruses and human influenza A virus. In particular, a part of the GP gene region of Ebola virus has been reported to encode two different proteins in different frames by transcriptional editing (Volchkov et al. 1995; Sanchez et al. 1996). Such a region could be influenced by a strong functional constraint. Although we have excluded that region from our analysis, we could not rule out the possibility that there are other overlapping regions currently unknown. At any rate, the relatively slow rate of nonsynonymous substitutions may be useful for establishing effective vaccines for these viruses, because the rate of emergence of a new phenotype as a source of the human infection may also be slow.

Divergence Times Among Ebola Virus Strains and Between Ebola and Marburg Viruses

The divergence times among Ebola virus strains and between Ebola and Marburg viruses were estimated on the assumption that these viruses had evolved at almost the same substitution rate. Ebola virus strains are known to be classified into four subtypes: Zaire, Sudan, Reston, and Ivory Coast subtypes (fig. 1). From the analysis of the rate of nonsynonymous substitutions for the GP gene of Ebola virus, we estimated that the Zaire and Ivory Coast subtypes diverged 700–1,300 years ago, the Sudan and Reston subtypes diverged 1,400–1,600 years ago, and these two clusters diverged 1,000–2,100 years ago. Moreover, the divergence time between Ebola and Marburg viruses was estimated to be 7,100–7,900 years ago. Similarly, from the analysis of the rate of nonsynonymous substitutions for Marburg virus, we estimated that these viruses diverged 1,000–9,800 years ago (table 3). Thus, although the divergence

Table 5
Relative Substitution Frequencies for the First and Second Codon Positions and Those for the Third Codon Position in the Entire Coding Region for Marburg Virus

Substitution Between	First and Second Positions [80.2]	Third Position [89.7]
A ↔ G	42.4 (87)	42.8 (220)
A ↔ T	3.0 (7)	1.8 (13)
A ↔ C	7.9 (17)	3.9 (20)
G ↔ T	5.4 (10)	4.2 (22)
G ↔ C	3.4 (6)	0.5 (2)
T ↔ C	37.9 (73)	46.9 (248)
Correlation Coefficient	−0.35	−0.25

NOTE.—The numbers in brackets represent proportions of transition substitutions. The numbers in parentheses represent raw numbers of nucleotide substitutions. Correlation coefficients between the frequencies of nucleotide substitutions and the chemical distances between two nucleotide bases, as defined by Gojobori, Li, and Graur (1982), are shown in the last row.

times estimated are in the wide range, we conclude that Ebola and Marburg viruses diverged more than several thousand years ago.

The divergence time between human immunodeficiency virus type 1 (HIV-1) and type 2 (HIV-2) has been estimated to be 150–200 years ago (Gojobori, Moriyama, and Yokoyama 1988). Moreover, hepatitis C virus has been estimated to have diverged from its ancestor around 300 years ago (Mizokami, Gojobori, and Lau 1994). In comparison with the estimated divergence times for these prevalent pathogenic viruses, Ebola and Marburg viruses might have diverged much earlier.

Patterns of Nucleotide Substitutions for Marburg Virus

The patterns of nucleotide substitutions at the first and second codon positions and the third codon position for Marburg virus were examined separately, as summarized in table 5. As most of the nucleotide substitutions at the first and second codon positions change coding amino acids, the substitution pattern at these positions would be influenced by natural selection at the protein level. On the other hand, as most of the substitutions at the third codon position do not change coding amino acids, nucleotide changes at that position are mostly free from natural selection (Kimura 1983) and reflect, to some extent, the pattern of spontaneous mutations in the genome.

At the third codon position in the entire coding region of Marburg virus, the proportion of transition substitutions was 90%, which was much larger than that of transversion substitutions. Furthermore, among transition substitutions, the frequencies of substitutions between purines were almost the same as those between pyrimidines. This feature of transition substitutions for Marburg virus is similar to that of influenza A virus (Saitou 1987). For HIV (Shimizu et al. 1989; Moriyama et al. 1991) and oncoviruses (Gojobori and Yokoyama 1987), however, substitution between purines is more frequent than that between pyrimidines. Although HIV and oncoviruses replicate themselves with reverse transcriptase, Marburg virus and influenza A virus have their own RNA-dependent RNA polymerase. Thus, the difference in transition substitutions among these viruses appears to reflect differences in the generating mechanisms of spontaneous mutations with viral polymerases.

We also investigated the pattern of nucleotide substitutions at the first and second codon positions of Marburg virus to examine whether any functional constraints are imposed on amino acid changes. For this purpose, we calculated the correlation coefficients between the frequencies of nucleotide substitutions at various codon positions and the chemical distances between two nucleotide bases, as defined by Gojobori, Li, and Graur (1982) (table 5). The chemical distance between two nucleotides was defined using Grantham's (1974) chemical distances between two amino acids. When a correlation coefficient is negative, it is possible that purifying selection may be operating on the nucleotide substitutions. Using this method, Gojobori, Li, and Graur (1982) demonstrated that purifying selection has operated on most of the eukaryotic functional genes. Saitou (1987) suggested that purifying selection has also operated on influenza A virus. For Marburg virus, the correlation coefficients were −0.35 for the first and second codon positions and −0.25 for the third codon position. Thus, it seems that the correlation coefficient for the first and second codon positions is larger than that for the third codon position, indicating that purifying selection has operated on Marburg virus during evolution.

This conclusion is supported by a recent study (Bukreyev et al. 1995b), in which 72.6% of the nucleotide substitutions in the entire coding region were found at the third codon position between the two strains of Marburg virus analyzed in this study. This is because purifying selection results in more frequent nucleotide substitutions at the third codon position than at the first and second codon positions (Kimura 1983). We also calculated the numbers of synonymous and nonsynonymous substitutions for the entire coding region between those two strains using the method of Nei and Gojobori (1986). The numbers of synonymous and nonsynonymous substitutions were estimated to be 0.180 ± 0.008 per site and 0.017 ± 0.001 per site, respectively. Thus, the number of synonymous substitutions was significantly higher than that of nonsynonymous ones ($P < 0.001$). This suggests that purifying selection has operated on Marburg virus during evolution, because synonymous substitutions are considered to be selectively neutral at the protein level, whereas nonsynonymous substitutions are influenced by selective constraints (Kimura 1983; Hughes and Nei 1988, 1989). The purifying selection would be caused by the functional constraint for viral proteins.

The pattern of nucleotide substitutions and the degree of functional constraints, which were estimated in the present study, are useful for the development of antiviral drugs and effective vaccines. In particular, inhibitors against viral replication will be able to be developed by taking into account the pattern of nucleotide substitutions. Moreover, if more data for the genome sequences of Ebola virus become available, it will be

possible to identify the gene product targeted by the host immune system, by comparing the degrees of functional constraints gene by gene. This is because the degree of functional constraints may vary with genes in the viral genome depending on the variability of amino acids.

In this study, we found that the GP gene of Ebola virus is evolving at the average rate of 3.6×10^{-5} per site per year at the nonsynonymous site, and that Marburg virus is evolving at a similar rate. These rates are of almost the same order of magnitude, but somewhat slow, in comparison with other RNA viruses. In particular, those rates are approximately a hundred times slower than those of retroviruses and human influenza A virus. We also estimated the divergence time between Ebola and Marburg viruses to be more than several thousand years ago. In addition, the pattern of nucleotide substitutions for Marburg virus indicated that the purifying selection has operated on this virus during evolution. These results will be useful in elucidating the origin and evolution of Ebola and Marburg viruses. To confirm and extend our observations, more sequence data for estimating the rate of synonymous substitutions and experimental studies on the mutation rates of Ebola and Marburg viruses would be required.

Acknowledgments

We thank Dr. Osamu Nakagomi at Akita University School of Medicine, Dr. Ian Humphery-Smith at Sydney University, and two anonymous reviewers for valuable suggestions. This study was supported, in part, by grants from the Ministry of Education, Science, Sports, and Culture.

LITERATURE CITED

BARON, R. C., J. B. MCCORMICK, and O. A. ZUBEIR. 1983. Ebola hemorrhagic fever in southern Sudan: hospital dissemination and intrafamilial spread. Bull. World Health Organ. **6**:997–1003.

BUKREYEV, A. A., E. F. BELANOV, V. M. BLINOV, and S. V. NETESOV. 1995*a*. Complete nucleotide sequences of Marburg virus genes 5 and 6 encoding VP30 and VP24 proteins. Biochem. Mol. Biol. Int. **35**:605–613.

BUKREYEV, A. A., V. E. VOLCHKOV, V. M. BLINOV, S. A. DRYGA, and S. V. NETESOV. 1995*b*. The complete nucleotide sequence of the Popp (1967) strain of Marburg virus: a comparison with the Musoke (1980) strain. Arch. Virol. **140**:1589–1600.

BUKREYEV, A. A., V. E. VOLCHKOV, V. M. BLINOV, and S. V. NETESOV. 1993*a*. The VP35 and VP40 proteins of filoviruses. Homology between Marburg and Ebola viruses. FEBS Lett. **322**:41–46.

———. 1993*b*. The GP-protein of Marburg virus contains the region similar to the 'immunosuppressive domain' of oncogenic retrovirus P15E proteins. FEBS Lett. **323**:183–187.

CENTERS FOR DISEASE CONTROL AND PREVENTION. 1995. Update: outbreak of Ebola viral hemorrhagic fever—Zaire, 1995. Morbid. Mortal. Wkly. Rep. **44**:468–475.

COX, N. J., J. B. MCCORMICK, K. M. JOHNSON, and M. P. KILEY. 1983. Evidence for two subtypes of Ebola virus based on oligonucleotide mapping of RNA. J. Infect. Dis. **147**:272–275.

FELDMANN, H., E. MUHLBERGER, A. RANDOLF, C. WILL, M. P. KILEY, A. SANCHEZ, and H.-D. KLENK. 1992. Marburg virus, a filovirus: messenger RNAs, gene order and regulatory elements of the replication cycle. Virus Res. **24**:1–19.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.

GOJOBORI, T., W.-H. LI, and D. GRAUR. 1982. Patterns of nucleotide substitution in pseudogenes and functional genes. J. Mol. Evol. **18**:360–369.

GOJOBORI, T., E. N. MORIYAMA, and M. KIMURA. 1990. Molecular clock of viral evolution and neutral theory. Proc. Natl. Acad. Sci. USA **87**:10015–10018.

GOJOBORI, T., E. N. MORIYAMA, and S. YOKOYAMA. 1988. Molecular evolutionary analysis of AIDS viruses. Pp. 142 *in* Fourth International Conference on AIDS Book 1. Stockholm.

GOJOBORI, T., Y. YAMAGUCHI, K. IKEO, and M. MIZOKAMI. 1994. Evolution of pathogenic viruses with special reference to the rates of synonymous and nonsynonymous substitutions. Jpn. J. Genet. **69**:481–488.

GOJOBORI, T., and S. YOKOYAMA. 1985. Rates of evolution of the retroviral oncogene of Moloney murine sarcoma virus and of its cellular homologues. Proc. Natl. Acad. Sci. USA **82**:4198–4201.

———. 1987. Molecular evolutionary rates of oncogenes. J. Mol. Evol. **26**:148–156.

GRANTHAM, R. 1974. Amino acid difference formula to help explain protein evolution. Science **185**:862–864.

HAYASHIDA, H., H. TOH, R. KIKUNO, and T. MIYATA. 1985. Evolution of influenza virus genes. Mol. Biol. Evol. **2**:289–303.

HUGHES, A. L., and M. NEI. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature **335**:167–170.

———. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc. Natl. Acad. Sci. USA **86**:958–962.

INA, Y., M. MIZOKAMI, K. OHBA, and T. GOJOBORI. 1994. Reduction of synonymous substitutions in the core protein gene of hepatitis C virus. J. Mol. Evol. **38**:50–56.

INTERNATIONAL COMMISSION. 1978. Ebola haemorrhagic fever in Zaire, 1976. Bull. World Health Organ. **56**:271–293.

KILEY, M. P., E. T. W. BOWEN, G. A. EDDY et al. (17 co-authors). 1982. Filoviridae: a taxonomic home for Marburg and Ebola viruses? Intervirology **18**:24–32.

KIMURA, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge.

LI, W.-H., C.-C. LUO, and C.-I. WU. 1985. Evolution of DNA sequences. Pp. 1–94 *in* R. J. MACINTYRE, ed. Molecular evolutionary genetics. Plenum Press, New York, N.Y.

LI, W.-H., M. TANIMURA, and P. M. SHARP. 1988. Rates and dates of divergence between AIDS virus nucleotide sequences. Mol. Biol. Evol. **5**:313–330.

MARTINI, G. A., and R. SIEGERT. 1971. Marburg virus disease. 1st edition. Springer, Berlin, Heidelberg, New York.

MIZOKAMI, M., T. GOJOBORI, and J. Y. N. LAU. 1994. Molecular evolutionary virology: its application to hepatitis C virus. Gastroenterology **107**:1181–1182.

MORIYAMA, E. N., Y. INA, K. IKEO, N. SHIMIZU, and T. GOJOBORI. 1991. Mutation pattern of human immunodeficiency virus genes. J. Mol. Evol. **32**:360–363.

MUEHLBERGER, E., A. SANCHEZ, A. RANDOLF, C. WILL, M. P. KILEY, H.-D. KLENK, and H. FELDMANN. 1992. The nucleotide sequence of the L gene of Marburg virus, a filovirus: homologies with paramyxoviruses and rhabdoviruses. Virology **187**:534–547.

NEI, M., and T. GOJOBORI. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous substitutions. Mol. Biol. Evol. **3**:418–426.

ORITO, E., M. MIZOKAMI, Y. INA, E. N. MORIYAMA, N. KAMESHIMA, M. YAMAMOTO, and T. GOJOBORI. 1989. Host-independent evolution and a genetic classification of the hepadnavirus family based on nucleotide sequences. Proc. Natl. Acad. Sci. USA **86**:7059–7062.

PRINGLE, C. R. 1991. The order Mononegavirales. Arch. Virol. **117**:137–140.

REGNERY, R. L., K. M. JOHNSON, and M. P. KILEY. 1981. Marburg and Ebola viruses: possible members of a new group of negative strand viruses. Pp. 971–977 *in* D. H. L. BISHOP and R. W. COMPANS, eds. The replication of negative strand viruses. Elsevier/North Holland, Amsterdam.

RICHARDSON, J. H., and W. E. BARKLEY. 1988. Biosafety in microbiological and biomedical laboratories. USPH, CDC. HHS Publication no. 88-8395.

SAITOU, N. 1987. Patterns of nucleotide substitutions in influenza A virus genes. Jpn. J. Genet. **62**:439–443.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SANCHEZ, A., M. P. KILEY, B. P. HOLLOWAY, and D. D. AUPERIN. 1993. Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. Virus Res. **29**:215–240.

SANCHEZ, A., M. P. KILEY, B. P. HOLLOWAY, J. B. MCCORMICK, and D. D. AUPERIN. 1989. The nucleoprotein gene of Ebola virus: cloning, sequencing, and in vitro expression. Virology **170**:81–91.

SANCHEZ, A., M. P. KILEY, H.-D. KLENK, and H. FELDMANN. 1992. Sequence analysis of the Marburg virus nucleoprotein gene: comparison to Ebola virus and other non-segmented negative-strand RNA viruses. J. Gen. Virol. **73**:347–357.

SANCHEZ, A., S. G. TRAPPIER, B. W. J. MAHY, C. J. PETERS, and S. T. NICHOL. 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.

SHIMIZU, N., T. OKAMOTO, E. N. MORIYAMA, Y. TAKEUCHI, T. GOJOBORI, and H. HOSHINO. 1989. Patterns of nucleotide substitutions and implications for the immunological diversity of human immunodeficiency virus. FEBS Lett. **250**: 591–595.

THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 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.

VOLCHKOV, V. E., S. BECKER, V. A. VOLCHKOVA, V. A. TERNOVOJ, A. N. KOTOV, S. V. NETESOV, and H.-D. KLENK. 1995. GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. Virology **214**:421–430.

WHO/INTERNATIONAL STUDY TEAM. 1978. Ebola haemorrhagic fever in Sudan, 1976. Bull. World Health Organ. **56**:247–270.

MITIKO GO, reviewing editor

Accepted April 25, 1997