Intragenic recombinations in rotaviruses

Yoshiyuki Suzuki[a,b], Takashi Gojobori[b,*], Osamu Nakagomi[a]

[a] *Department of Microbiology, Akita University School of Medicine, Akita, Japan*
[b] *Center for Information Biology, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka-ken, 411-8540, Japan*

Received 23 March 1998; revised version received 3 April 1998

Abstract In this paper, evidence for intragenic recombination in the VP7 gene between rotavirus strains bearing different serotypes is demonstrated for the first time. Intragenic recombination may be one of the escaping mechanisms from the host immune system for rotavirus. This process involves exchanging antigenic regions, thus questioning the use of multivalent vaccines for the prevention of rotavirus infection.

Key words: Intragenic recombination; Evolution; Vaccine; Rotavirus

1. Introduction

Group A–F rotaviruses are members of the genus *Rotavirus* in the family Reoviridae. Among these viral groups, groups A, B, and C are known to infect humans, and, moreover, group A rotavirus is one of the most important etiological agents of gastroenteritis in infants worldwide. The virion of this virus is a non-enveloped, triple-layered capsid containing the 11 segmented, double-stranded RNA genome [1,2].

The rotavirus genome encodes six structural (VP1–VP4, VP6, and VP7) and five non-structural (NSP1–NSP5) proteins. The structural proteins VP4 and VP7 constitute the outer shell of the virion. These two proteins are known to be involved in the neutralization of the viral particle, where the predominant reactivity is directed against VP7 [1,2].

In the case of VP7, 14 serotypes (designated G1–G14, after glycoprotein) have been distinguished, based on the criterion of a greater than 20-fold difference between the homologous and heterologous reciprocal neutralization antibody titers [3]. This protein is encoded by the seventh, eighth, or ninth gene segment of the genome. It contains nine regions (VR1–VR9) which are highly divergent among different serotypes but conserved within the same serotype [4]. By investigating the nucleotide sequence of escape mutants from monoclonal antibodies, it has been shown that VR5, VR7, and VR8 are the most important for the neutralization of rotavirus particles [5–8].

Nucleotide substitution and gene reassortment have been reported to be the major mechanisms of producing antigenic variants for rotavirus strains [9,10]; in particular, recombination has been reported to occur in an 'all or none' fashion [11]. Intragenic recombination has not been considered for rotavirus, though it has been established for many other RNA viruses [12,13].

In this paper, evidence of intragenic recombination for rotavirus is reported for the first time, by investigating VP7 gene sequences. Intragenic recombination may be one of the important mechanisms for rotavirus to produce escape mutants from the host immune system.

2. Materials and methods

2.1. Sequence data

Two hundred and seven sequence data for the VP7 gene of group A rotavirus, which did not include any gaps in the alignment for the entire coding region, were collected from the international DNA databanks (DDBJ/EMBL/GenBank).

2.2. Data analysis

The sequence alignment was made by CLUSTAL W [14]. The phylogenetic tree was reconstructed by the neighbor-joining method [15] with the evolutionary distance estimated by Kimura's two-parameter method [16]. The bootstrap probability was calculated for each internal branch of the phylogenetic tree with 10 000 times of resampling [17].

The border of the recombination was identified by the method of Robertson et al. [18] with some modifications. In this method, four sequences were used for analysis: one putative recombinant sequence, one representative sequence from each of the two parental lineages considered to be involved in the recombination, and one sequence from known outgroup. An alignment was made of the four sequences, and the informative sites were collected along the alignment. The informative site supported either of the three phylogenetic relationships: the clustering of the recombinant with one of the two parental sequences, with the other parental sequence, and with the outgroup sequence. The relative frequencies of the former two types of sites were compared between the regions before and after each breakpoint along the alignment, using the chi-square value. If the value was statistically significant at some breakpoints, the optimum position of the recombination was determined as the point which maximizes that value [18].

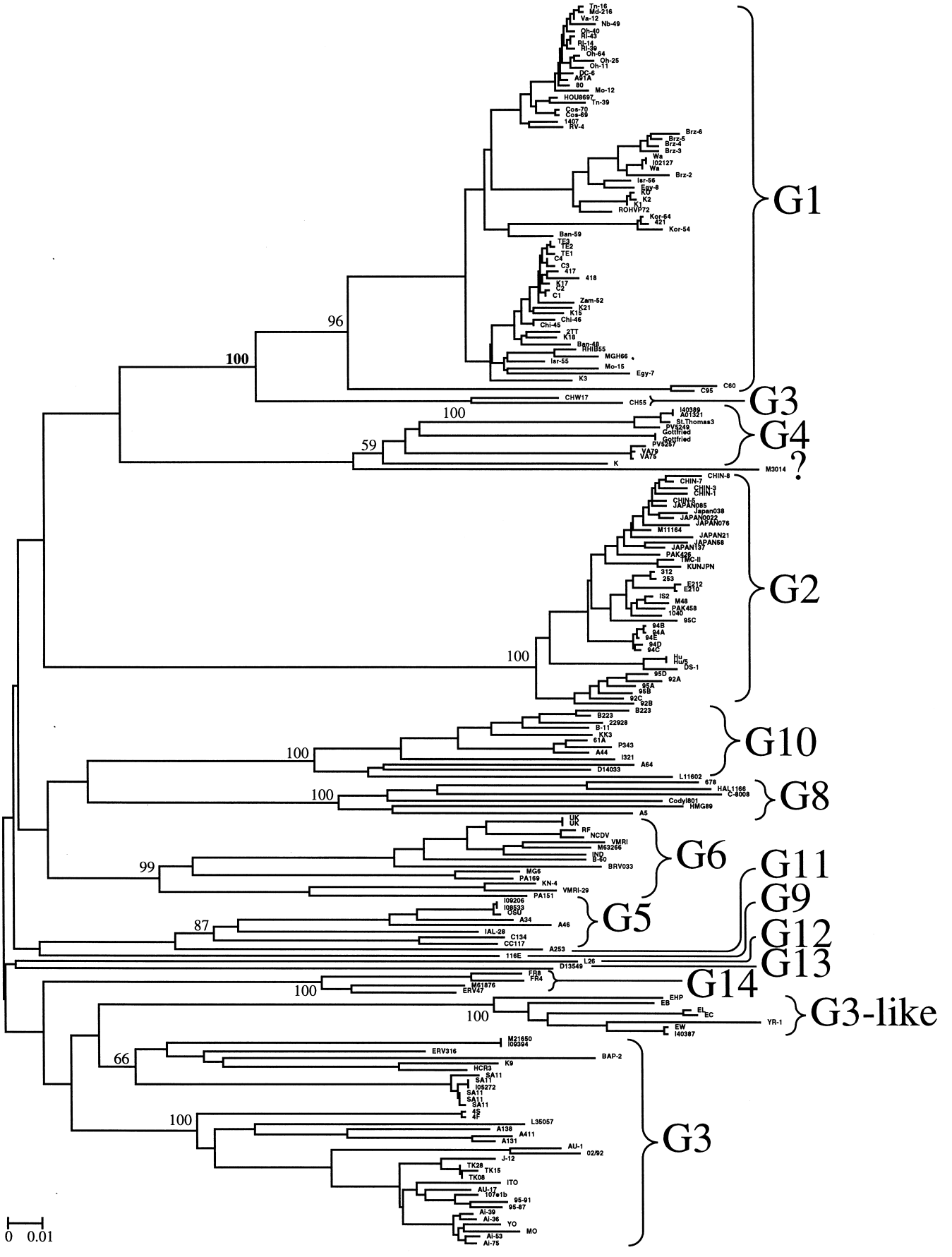
In the present study, a sequence which was considered to be unrelated to the recombination event was used instead of the outgroup sequence, because the outgroup was not available for use. However, it was considered that the result should not be affected by this modification, because we were concerned only with the distribution of the informative sites. Moreover, Fisher's exact test was used instead of the chi-square test to detect significance, because the actual numbers in 2×2 contingency tables were relatively small [19].

3. Results

When the phylogenetic tree was reconstructed for all sequence data using the entire coding region of the VP7 gene, it was found that each of the serotypes formed a single cluster mostly with high bootstrap probability (Fig. 1). However, there were two peculiar strains which did not follow the general tendency observed above. In fact, CH55 and CHW17, which were originally determined as G3 by serological study using monoclonal antibodies [20], did not make a single cluster with other G3 strains, but made a cluster with G1 strains with a bootstrap probability of 100% (Fig. 1). There are two possible mechanisms which could explain that observation: intragenic recombination and sequence convergence.

*Corresponding author. Fax: (81) (559) 81 6848.
E-mail: tgojobor@genes.nig.ac.jp

PII S0014-5793(98)00415-3


Tn-16
Md-216
Va-12
Nb-49
Oh-40
RI-43
RI-14
RI-36
OH-64
Oh-25
Oh-11
DC-6
A81A
80
Mo-12
HOU8697
Tn-39
Cos-70
Cos-69
1407
RV-4
Brz-6
Brz-5
Brz-4
Brz-3
Wa
I02127
Wa
Brz-2
Isr-56
Egy-8
KU
K2
V3
ROHVP72
Kor-54
421
Kor-54
Ban-59
TE3
TE2
TE1
C4
C3
417
418
K17
C2
C1
Zam-52
K21
K15
Chi-46
Chi-45
2TT
K18
Ban-48
RHIB55
MGH66
Isr-55
Mo-15
Egy-7
K3
C60
C95
CHW17
CH55
I40389
A01321
St.Thomas3
PV5249
Gottfried
Gottfried
PV5257
VA79
VA75
K
M3014
CHIN-8
CHIN-7
CHIN-3
CHIN-1
CHIN-5
JAPAN085
Japan038
JAPAN022
JAPAN076
M11164
JAPAN21
JAPAN58
JAPAN137
PAK426
TMC-II
KUNJPN
312
253
E212
E210
IS2
M49
PAK458
1040
95C
94B
94A
94E
94D
94C
Hu/s
DS-1
95D
92A
95A
95B
92C
92B
B223
B223
22928
B-11
KK3
61A
P343
A44
I321
A64
D14033
L11602
678
HAL1166
C-8008
Cody801
HMG89
A5
UK
UK
RF
NCDV
VMRI
M63266
IND
B-60
BRV033
MG6
PA169
KN-4
VMRI-29
PA151
I09206
I08533
OSU
A34
A46
IAL-28
C134
CC117
A253
116E
D13549
L26
FR8
FR4
M61876
ERV47
EHP
EB
EL
EC
YR-1
EW
I40387
M21650
I09394
ERV316
BAP-2
K9
HCR3
SA11
SA11
I05272
SA11
SA11
SA11
4S
4F
L35057
A138
A411
A131
AU-1
02/92
J-12
TK28
TK15
TK08
ITO
AU-17
107e1b
95-91
95-87
AI-39
AI-36
YO
MO
AI-53
AI-75
96
100
100
59
100
100
100
99
87
100
100
66
100
G1
G3
G4
?
G2
G10
G8
G6
G11
G9
G5
G12
G13
G14
G3-like
G3
0 0.01

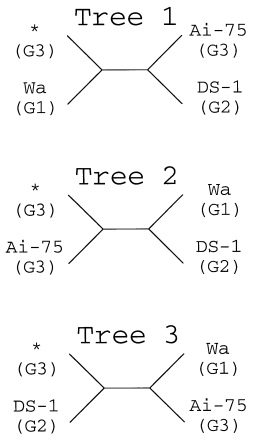


Fig. 2. The three possible phylogenetic relationships supported by the informative site in the alignment of Wa, DS-1, Ai-75, and CH55 or CHW17. * indicates CH55 or CHW17.

To examine whether intragenic recombination took place during the evolution of CH55 and CHW17, a sequence comparison for viral isolates was made. In practice, four sequences were chosen in order to conduct the method of Robertson et al. [18]. They were CH55 or CHW17, one representative from each of the G1 and G3 strains which were considered to be involved in the recombination event, and one sequence from the G2 strains which were considered to be unrelated to the recombination. Wa, DS-1, and Ai-75 were used as the representative sequences for G1, G2, and G3 strains, respectively. Nucleotide sequences for the four strains were aligned, and informative sites collected as shown in Table 1. These sites were classified into three groups, according to the phylogenetic relationships supported (Fig. 2).

From Table 1, it was hypothesized that each of the two sequences could be divided into three subgenic regions, namely regions A, B, and C for CH55, and regions a, b, and c for CHW17, as shown in Fig. 3. This hypothesis was proven by statistical analysis in which boundaries were detected with high statistical significance ($P = 9.2\times10^{-10}$ and 5.9×10^{-8} between regions A and B, and B and C, respectively, and 1.1×10^{-3} and 1.1×10^{-7} between regions a and b, and b and c, respectively). When a phylogenetic tree was made for each of these regions using the four sequences compared, regions A, C, a, and c supported topology 2 (Fig. 2), and regions B and b supported topology 1 (Fig. 2), all with significantly high bootstrap probabilities (Fig. 3). Additional boundaries were detected in some regions: region B could be divided into B′ and B″, region a into a′ and a″, and region b into b′ and b″, all with relatively low statistical significance ($(2.0\text{–}4.3)\times10^{-2}$) (Fig. 3). However, the phylogenetic tree reconstructed for each of these subdivided regions supported the same topology as the original one (Fig. 3).

Moreover, a phylogenetic tree was reconstructed for each of these divided regions (a′, a″, B′(b′), B″(b″), and C(c)) using 207 sequence data (data not shown). For all regions except a′, the topology of the phylogenetic tree supported that CH55 and CHW17 belonged to the same serotype as was indicated in the four sequence analysis. For region a′, however, CHW17 was closer to G1 than any other serotypes, which was inconsistent with the above result.

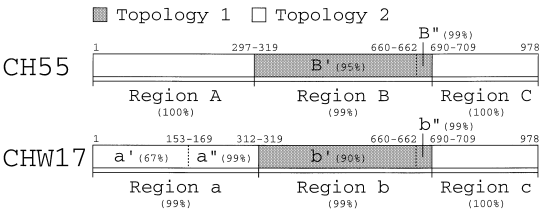


Fig. 3. The mosaic structure in the coding region of the VP7 gene for CH55 and CHW17. The topology supported by each region is indicated by color, with bootstrap probability in parentheses. The topology is described in Fig. 2.

From these results, it was concluded that cross-serotypic intragenic recombination had occurred between sequences G1 and G3 with crossover points at nucleotide site 297–319 and 690–709 of CH55, and 312–319 and 690–709 of CHW17 (Fig. 3). An additional recombination between different G1 strains may have occurred at site 660–662 of CH55 and CHW17. Because the above three crossover sites were almost the same for both strains, these recombination events may have occurred in the ancestor of CH55 and CHW17. Although another recombination may have occurred at site 153–169 only on the lineage for CHW17, it could not be determined which serotype region a′ was most closely related, due to the low bootstrap probability in the phylogenetic tree (Fig. 3).

4. Discussion

It was concluded that the mosaic pattern of nucleotide sequences observed in CH55 and CHW17 was derived from the recombination event, rather than sequence convergence. If it was assumed that the mosaic pattern was caused by sequence convergence due to positive selection, informative sites should be observed mainly at the first and second codon positions because selective force operates mainly at the amino acid sequence level [21]. However, most of the informative sites were located at the third codon position (Table 1). Moreover, if the mosaic pattern was assumed to be derived from the convergence by random fluctuation of sequence variation, the informative sites supporting a particular phylogenetic relationship should be evenly scattered throughout the sequence, like the sites supporting topology 3 (Fig. 2, Table 1). This was not the case for the sites supporting topologies 1 and 2. As a result, recombination was much more likely to be the cause of the mosaic pattern than sequence convergence.

Recombination should not occur unless multiple rotavirus strains infect the same cell. Indeed, multiple infection has long been reported for rotavirus, particularly through the existence of many reassortant strains in nature [9,10]. Thus, there should be significant opportunities for recombination in the evolutionary pathway of rotavirus. However, the molecular

←
Fig. 1. The phylogenetic tree for 207 sequences of the VP7 gene for group A rotavirus, using the entire coding region. The serotype (G type) is described on the right side of the tree. The bootstrap probability is indicated on the branch which supports the clustering of strains with the same serotype.

Table 1
Informative sites in the alignment of Wa, DS-1, Ai-75, and CH55 or CHW17, and the topology supported by them (o)

Site	CH55			CHW17		
	Topology			Topology		
	1	2	3	1	2	3
25		o			o	
28		o				
48		o			o	
52				o		
55				o		
63		o			o	
64		o			o	
67		o			o	
69		o			o	
78				o		
95		o				
96		o				
102				o		
105				o		
112		o			o	
123		o				
126				o		
132		o			o	
135				o		
143				o		
144		o			o	
146		o			o	
153				o		
169		o			o	
171		o			o	
174		o				
177		o			o	
193		o			o	
195		o				
196		o			o	
213		o			o	
219		o			o	
224						o
225	o					
228		o			o	
240		o				
249			o			o
255		o			o	
268		o			o	
274				o		
289		o			o	
290		o			o	
297		o			o	
312					o	
319	o			o		
324	o					
330	o					
364	o			o		
366	o			o		
367						o
375				o		
377				o		
390	o					
397		o				
399		o			o	
402		o			o	
421	o			o		
437					o	
438	o			o		
441	o			o		
448					o	
462		o			o	
468	o			o		

Table 1 (*continued*)
Informative sites in the alignment of Wa, DS-1, Ai-75, and CH55 or CHW17, and the topology supported by them (o)

Site	CH55			CHW17		
	Topology			Topology		
	1	2	3	1	2	3
472	o			o		
478					o	
501	o			o		
519		o			o	
532				o		
534	o			o		
536		o				
549	o			o		
558		o			o	
577		o				
579	o			o		
612		o			o	
615	o			o		
621				o		
627				o		
631	o					
633		o			o	
635					o	
637	o			o		
640	o			o		
649					o	
659	o					
660		o			o	
662	o			o		
666			o			o
674	o			o		
678	o			o		
679	o			o		
680	o			o		
681	o			o		
684	o			o		
687	o			o		
690	o			o		
709		o			o	
711		o			o	
714					o	
720		o			o	
735		o			o	
739		o			o	
765		o			o	
771		o			o	
777		o			o	
780		o			o	
786		o			o	
799		o			o	
801			o			o
805		o			o	
807		o			o	
828		o			o	
832	o			o		
843		o			o	
870		o			o	
906		o			o	
907		o			o	
909		o			o	
921		o			o	
926		o			o	
942		o			o	
963		o			o	
969		o			o	
972		o			o	
975		o			o	
976	o			o		

The topology is described in Fig. 2.

mechanisms of the recombination is still unclear. For other RNA viruses, homologous recombination and template switching have been reported as the mechanisms of intragenic recombination, and the latter has been reported as the main mechanism in many cases [22]. For rotavirus, however, it is known that the synthesis of both RNA strands of the genome is particle-associated, and free double-stranded or minus-stranded RNA has never been found in infected cells [1]. This fact seems to reduce the probability of both homologous recombination and template switching in the replication process of this virus.

Intragenic recombination may be one of the mechanisms of evolution for rotavirus, especially by producing escape mutants through exchanging antigenic regions between strains with different antigenicity. By investigating the nucleotide sequence of escape mutants from monoclonal antibodies, it has been reported that VR5, VR7, and VR8 are the most important for the neutralization of rotavirus particles [5–8]. The identified crossover points for inter-serotypic recombination in CH55 and CHW17 were located between VR5 and VR6 and between VR8 and VR9, resulting in the virus strains possessing VR1–VR5 and VR9 related to G3, and VR6–VR8 related to G1 (Fig. 3). An experimental study has shown that the monoclonal antibody against serotype 3 reacted with CH55, whereas that against serotype 1 failed [20]. If the monoclonal antibodies used in that experiment recognized VR5 of CH55, it is consistent with the idea of intragenic recombination.

The currently predominant approach for prevention from rotavirus infection is to inoculate the multivalent reassortant vaccine covering serotypes prevalent in the community [23]. However, it has been reported that the mutants possessing a mutation at different sites on the VP7 gene could escape from a particular monoclonal antibody, indicating that the interaction among different sites is involved in the antigenicity of VP7 [5–8]. If this is the case, it is consistent with the idea that intragenic recombination produces new variants of rotavirus which escape from the immune response elicited by multivalent vaccines, by constructing mosaic structures in the VP7 and VP4 genes.

The frequency of intragenic recombination for rotavirus may be low in nature, because almost all isolates were divided into distinct clusters according to their serotypes in the phylogenetic tree (Fig. 1), and there seems to be little chance for recombination in the replication process of rotavirus. However, the intra-serotypic recombination, which is more difficult to detect, may be more frequent than the inter-serotypic recombination, as is the case for intra- and inter-genogroup reassortment for this virus [9]. Moreover, even if the frequency of intragenic recombination were low, it should still be an important mechanism of evolution for rotavirus, because it can create variants with dramatic changes in their antigenicity.

Acknowledgements: We are grateful to Prof. Hiroshi Ushijima at Tokyo University, Japan, for providing sequence data. We thank Dr. Silvana Gaudieri at National Institute of Genetics, Japan, for valuable comments on the manuscript. This study was supported, in part, by the Ministry of Education, Science, Sport, and Culture of Japan.

References

[1] Estes, M.K. (1996) in: Fields Virology (Fields, B.N., Knipe, D.M., Howley, P.M., Chanock, R.M., Melnick, J.L., Monath, T.P., Roizman, B. and Straus, S.E., Eds.). Vol. 2, 3rd edn., pp. 1625–1655, Lippincott-Raven, Philadelphia, PA.
[2] Kapikian, A.Z. and Chanock, R.M. (1996) in: Fields Virology (Fields, B.N., Knipe, D.M., Howley, P.M., Chanock, R.M., Melnick, J.L., Monath, T.P., Roizman, B. and Straus, S.E., Eds.), Vol. 2, 3rd edn., pp. 1657–1708, Lippincott-Raven, Philadelphia, PA.
[3] Browning, C.G., Fitzgerald, T.A., Chalmers, R.M. and Snodgrass, D.R. (1991) J. Clin. Microbiol. 29, 2043–2046.
[4] Green, K.Y., Hoshino, Y. and Ikegami, N. (1989) Virology 168, 429–433.
[5] Dyall-Smith, M.L., Lazdins, I., Tregear, G.W. and Holmes, I.H. (1986) Proc. Natl. Acad. Sci. USA 83, 3465–3468.
[6] Mackow, E.R., Shaw, R.D., Matsui, S.M., Vo, P.T., Benfield, D.A. and Greenberg, H.B. (1988) Virology 165, 511–517.
[7] Taniguchi, K., Hoshino, Y., Nishikawa, K., Green, K.Y., Maloy, W.L., Morita, Y., Urasawa, S., Kapikian, A.Z., Chanock, R.M. and Gorziglia, M. (1988) J. Virol. 62, 1870–1874.
[8] Hoshino, Y. and Kapikian, A.Z. (1994) Curr. Top. Microbiol. Immunol. 185, 179–227.
[9] Nakagomi, O. and Nakagomi, T. (1993) Microbiol. Immunol. 37, 337–348.
[10] Nakagomi, O. and Nakagomi, T. (1996) Arch. Virol. 12, 93–98.
[11] Ramig, R.F. (1982) Virology 120, 92–105.
[12] Gojobori, T., Moriyama, E.N., Ina, Y., Ikeo, K., Miura, T., Tsujimoto, H., Hayami, M. and Yokoyama, S. (1990) Proc. Natl. Acad. Sci. USA 87, 4108–4111.
[13] Robertson, D.L., Sharp, P.M., McCutchan, F.E. and Hahn, B.H. (1995) Nature 374, 124–126.
[14] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Nucleic Acids Res. 22, 4673–4680.
[15] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.
[16] Kimura, M. (1980) J. Mol. Evol. 16, 111–120.
[17] Felsenstein, J. (1985) Evolution 39, 783–791.
[18] Robertson, D.L., Hahn, B.H. and Sharp, P.M. (1995) J. Mol. Evol. 40, 249–259.
[19] Sokal, R.R. and Rohlf, F.J. (1995) Biometry, 3rd edn., W.H. Freeman and Company, New York.
[20] Wen, L., Nakayama, M., Yamanishi, Y., Nishio, O., Fang, Z.-Y., Nakagomi, O., Araki, K., Nishimura, S., Hasegawa, A., Muller, W.E.G. and Ushijima, H. (1997) Arch. Virol. 142, 1481–1489.
[21] Kimura, M. (1983) The Neutral Theory of Molecular Evolution, Cambridge University Press, Cambridge.
[22] Kirkegaard, K. and Baltimore, D. (1986) Cell 47, 433–443.
[23] Bernstein, D.I., Glass, R.I., Rodger, G., Davidson, B.L. and Sack, D.A. (1995) J. Am. Med. Assoc. 273, 1191–1196.