



## Nucleotide mismatches of foot-and-mouth disease virus during replication

Y.B. Ma, C.X. Hao and H.Y. Chang

State Key Laboratory of Veterinary Etiological Biology,  
National Foot and Mouth Disease Reference Laboratory,  
Lanzhou Veterinary Research Institute,  
Chinese Academy of Agricultural Sciences, Lanzhou, China

Corresponding author: H.Y. Chang  
E-mail: changhuiyun@126.com

Genet. Mol. Res. 12 (2): 1022-1027 (2013)  
Received July 3, 2012  
Accepted August 16, 2012  
Published April 2, 2013  
DOI <http://dx.doi.org/10.4238/2013.April.2.18>

**ABSTRACT.** As there is a lack of error correction mechanisms during RNA replication, foot-and-mouth disease virus (FMDV) has a very high mismatch rate, which leads to a high mutation rate, in the range of  $10^{-3}$  to  $10^{-5}$  per nucleotide site per genome replication. We examined the nucleotide mismatch of FMDV during replication, based on the whole genomes of the 7 serotypes retrieved from NCBI. With the Mega bio-software, SPSS, and Microsoft Excel, we studied the nucleotide differences compared to the sequence in the RefSeq database, and developed two probable mutation models, i.e., once mutation model and complication mutation model. Further analysis on the nucleotide mismatch during replication was made. We found that FMDV share similar difference rates between nucleotide and reverse differences, for example the mutation  $U \rightarrow C$  and  $C \rightarrow U$ . We also found that each nucleotide has its domain mismatch, and the virus kept a constant nucleotide composition during mutations.

**Key words:** FMDV; Mismatch; Mutation; Nucleotide

## INTRODUCTION

Foot-and-mouth disease is one of the most contagious viral diseases of cloven-hoofed livestock including cattle, swine, sheep, goats, and over 70 species of wild animals. The disease results in high morbidity, lost productivity, high mortality in young animals, and low mortality in adult animals. Foot-and-mouth disease virus (FMDV), the first demonstrated filterable agent causing disease in animals, belongs to the *Aphthovirus* genus of the Picornaviridae family and contains a single-stranded positive-sense RNA genome of about 8500 nucleotides. FMDV lacks an error correction mechanism for RNA replication, and thus has a very high mismatch rate, which leads to mutation rates of  $10^{-3}$  to  $10^{-5}$  per nucleotide site per genome (Drake and Holland, 1999; Grubman and Baxt, 2004). This high mutation rate has produced 7 serotypes, multiple subtypes, and variants.

To understand the nucleotide mismatching that occurs during replication, we retrieved the genomes of the 7 FMDV serotypes from NCBI. We statistically analyzed the nucleotide differences between these sequences and the reference sequence in the RefSeq database, and generated 2 probable mutation models. Further analysis of nucleotide mismatch during replication revealed that the FMDV nucleotide and reverse difference rates are similar. In addition, each nucleotide has its domain mismatch, and the virus maintains a nucleotide composition similar to that of the mother virus during evolution.

## MATERIAL AND METHODS

### Retrieval of whole genome sequences and the processing of raw statistics

The whole genome sequences of FMDV were retrieved from NCBI. Analysis was conducted in 7 separate groups based on serotype. The genome sequences retrieved from NCBI were multiply aligned with Muscle in Mega. Then, directional nucleotide pair frequencies in Mega were used to analyze nucleotide differences between each sequence and the standard sequence in the RefSeq database, and the results were exported to Excel. Then, the mutation numbers were counted. As random mutation can only be defined after sufficient mutations have occurred, the cut-off number of nucleotide differences was set as 500. The directional difference frequency rate of each genome sequence was calculated and represented as a decimal. All statistical processing was conducted in SPSS 17.0 (Kerr et al., 2002).

### Construction of possible nucleotide mutation models

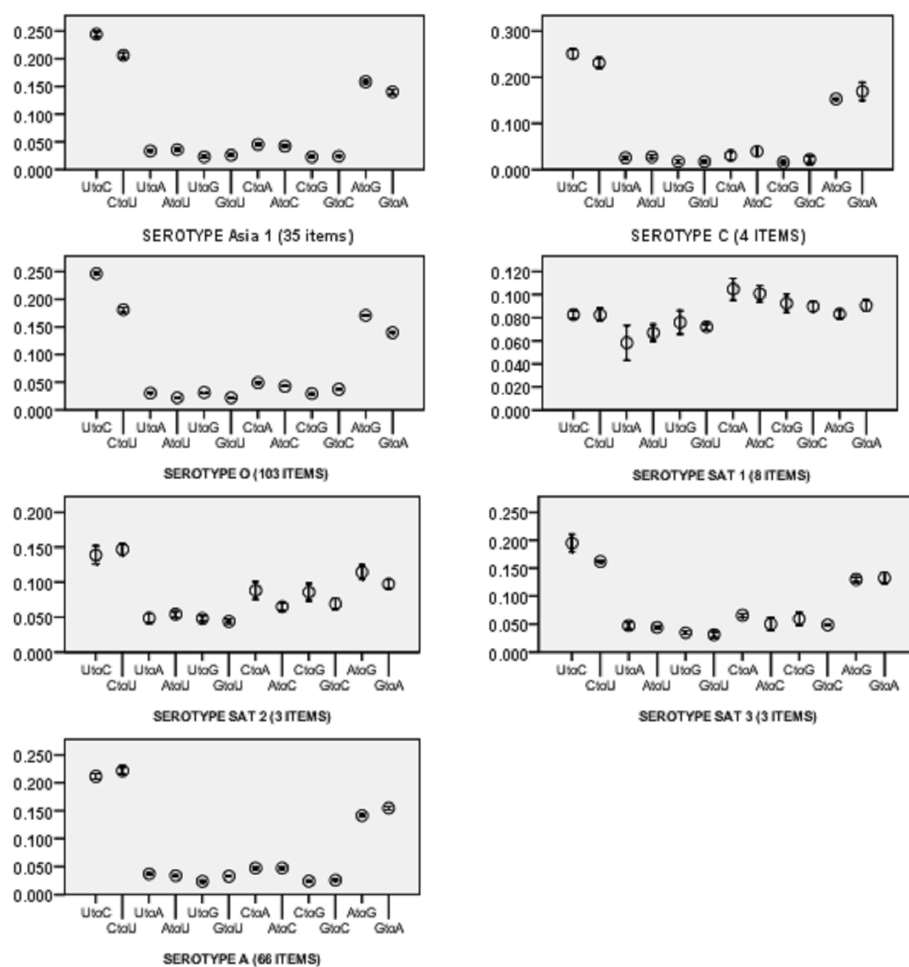
Construction of mutation models was based on the FMDV replication cycle and nucleotide base pair. FMDV replication proceeds from sense chain to antisense chain, then from antisense to sense (Belsham, 2005). Thus, mismatches may occur in either of 2 steps during a single-replication cycle.

## RESULTS

### The rates of nucleotide difference by serotype

At last, 103 items of serotype O, 66 items of serotype A, 4 items of serotype C, 35

items of serotype Asia 1, 8 items of SAT 1, 3 items of SAT 2, and 3 items of SAT 3 met the requirements. Selected sequences within each serotype were analyzed in a group. The 95% error nucleotide difference rate bars in each serotype are shown in Figure 1.



**Figure 1.** Nucleotide difference rates of the 7 serotypes. The 95% error nucleotide difference rate bars of the 7 serotypes are shown. The horizontal axis is the nucleotide difference, and the vertical axis is the difference rate of a nucleotide difference in all the nucleotide differences. For example, the U→G means the U is in standard sequence while G is in the same position of the compared sequence. The mean rate of U→G is 0.0311 in serotype O.

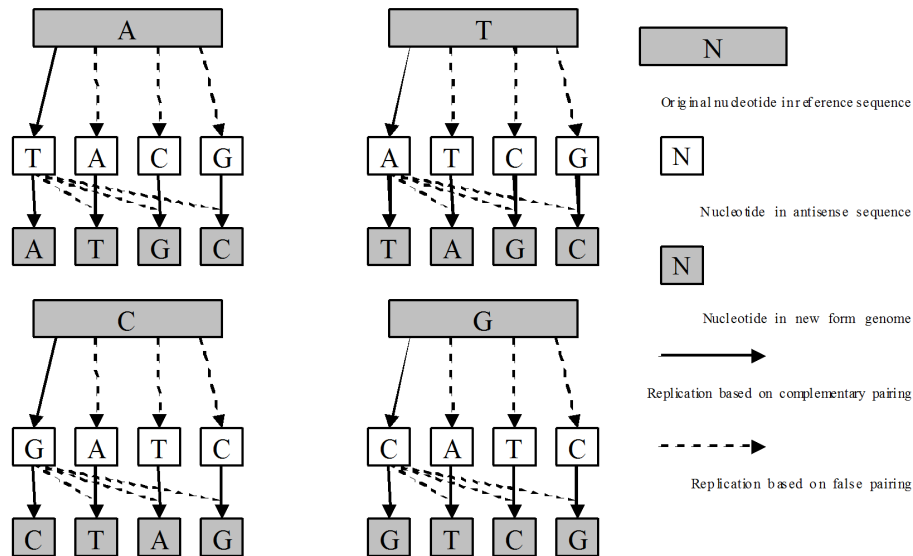
Furthermore, to confirm the nucleotide difference rates, the sequences were compared against a different reference sequence. Two random comparisons from each serotype were chosen, and their nucleotide difference rates were compared to the difference rates obtained from the RefSeq database (Beard and Mason, 2000; Carrillo et al., 2005). All P values were >0.01 (Morgan, 2007), indicating no significant difference in nucleotide difference rates in comparison to a non-reference standard.

### A possible nucleotide mutation model

Nucleotide differences between 2 different strains were caused by poor fidelity during RNA replication. FMDV replication proceeds from sense chain to antisense chain, then from antisense to sense. Numerous copies of complete viral genomes were synthesized by this process. Three nucleotide mutation models may account for the nucleotide differences caused by mutation.

#### Once mutation model

The once mutation model describes the case in which a nucleotide mutation in a unique position appears only once during FMDV replication. This model is shown in Figure 2.

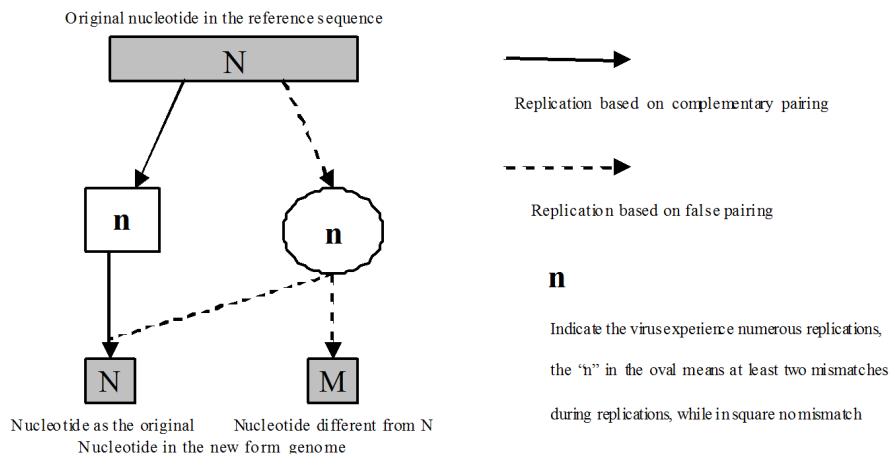


**Figure 2.** Once mutation model. Possible nucleotide difference that is lead by the once nucleotide mismatch during replication. “N” in rectangular shadow box represents the original nucleotide in the standard sequence, while “N” in square shadow box means the nucleotide in the same position of the newly formed genome. The “N” in the open square means the possible mismatch during replication. The dashed line is the mismatch while the solid one is the match based on base pair.

In this model, the nucleotide mutation has an equal chance of occurring in the sense-to-antisense step or the antisense-to-sense step. By this mechanism, the nucleotide may be changed in the newly formed virus genome.

#### Complication mutation model

The complication mutation model describes a case in which the difference between 2 sequences is the result of accumulated mutations over numerous generations of replication. This model is presented in Figure 3.



**Figure 3.** Complication mutation model. Over two times mismatch lead to a nucleotide difference. The meaning of symbols is shown in the figure.

The complication mutation model may yield 2 or more mismatches. When 2 mutations lead to the nucleotide difference, 2 circumstances existed. They were 2 mutations in the same cycle and 2 mutations in different cycles. In other cases, more than 2 mutations lead to a nucleotide difference.

## DISCUSSION

Mismatches during replication generated these differences. However, during replication of the FMDV genome in cells without reverse transcription, mismatches can occur in the sense-to-antisense step or the antisense-to-sense chain step. The 2-mutation models are theoretical. However, the once mutation model is likely significant. The FMDV mutation rate is  $10^{-3}$  to  $10^{-5}$  per nucleotide site per genome replication; thus, the mutation rate of a specific nucleotide is at most  $10^{-3}$ . Nucleotide mismatches occurred at random, and was thus an independent event from generation to generation. The possibility of 2-random mismatches occurring at a unique position during a single-replication cycle is at most  $10^{-3}$  (Parzen, 1960). The more the generation of nucleotide mismatch needed for the specific nucleotide difference, the smaller the possibility. This is nearly impossible even though some mismatches may occur in the same position over numerous generations. Based on this analysis, we conclude that the once mutation model is the main mechanism for nucleotide differences between sequences.

From the once mutation model, we can conclude that the U→C change from the reference sequence was caused by mismatches of U→G and A→C. Other relationships between nucleotide differences and mutations are described in Table 1.

It is interesting to note that all FMDV share similar difference rates between nucleotide and reverse differences. As the difference rates are similar, we can treat a nucleotide difference and its reverse difference as a difference group. It was obvious that FMDV share a similar nucleotide difference rate with the exception of serotype SAT. Groups U→C and A→G were the two most frequent difference groups; all others were at a similar low level. In serotype SAT, especially SAT 1, the 4 low difference groups had dissimilar difference rates.

Although the 4 low difference groups in serotype SAT significantly differed from the other serotypes, the 2 highest difference groups were the same, except in SAT 1. It was interesting to note that the nucleotide difference in SAT 1 seemed truly random. Further effort will be required to determine whether the exception in serotype SAT is due to geographical differences, small samples, or some other cause.

**Table 1.** Corresponding relationships between nucleotide difference and nucleotide mismatch.

Nucleotide mismatch	Nucleotide difference
U→G, A→C	U→C, A→G
U→U, A→A	U→A, A→T
G→A, C→U	G→U, C→A
G→U, C→A	C→U, G→A
C→C, G→G	C→G, G→C
U→C, A→G	A→C, U→G

The two different nucleotide mismatches can lead to the same nucleotide difference. These two different nucleotide mismatches can lead to different nucleotide difference. For example, the mismatches of U→G, A→C can not only lead to the difference of U→C, but also A→G.

Based on the nucleotide differences and Table 1, we can see that the most probable nucleotide mismatches during replication were U→G, A→C, G→U, and C→A. Each nucleotide has its domain mismatch nucleotide, and all were different. From the domain mismatch, we can also see that pyrimidine→pyrimidine, purine→purine, U→C, A→G, G→A, and C→U mismatches were extremely unlikely, leading to a low nucleotide difference rate in the 4 low groups.

It is intriguing that this mismatch mechanism allows the virus to maintain a nucleotide composition similar to that of the mother virus during its evolution. The above analysis revealed the existence of domain mismatches. Is this a result of poor RNA polymerase fidelity or of natural selection on the viral mutation? Further research may help us understand the molecular evolution of this virus and allow us to predict new strains to control the FMDV epidemics.

## ACKNOWLEDGMENTS

Research supported by the fund of Key Special Projects of Breeding New Varieties of Genetically Engineered Organisms (#2011ZX08011-004, #2009ZX08007-008B, #2009ZX08006-002B) from the Ministry of Agriculture of the People's Republic of China.

## REFERENCES

- Beard CW and Mason PW (2000). Genetic determinants of altered virulence of Taiwanese foot-and-mouth disease virus. *J. Virol.* 74: 987-991.
- Belsham GJ (2005). Translation and replication of FMDV RNA. *Curr. Top. Microbiol. Immunol.* 288: 43-70.
- Carrillo C, Tulman ER, Delhon G, Lu Z, et al. (2005). Comparative genomics of foot-and-mouth disease virus. *J. Virol.* 79: 6487-6504.
- Drake JW and Holland JJ (1999). Mutation rates among RNA viruses. *Proc. Natl. Acad. Sci. U. S. A.* 96: 13910-13913.
- Grubman MJ and Baxt B (2004). Foot-and-mouth disease. *Clin. Microbiol. Rev.* 17: 465-493.
- Kerr AW, Hall HK and Kozub S (2002). *Doing Statistics with SPSS*. vii. Thousand Oaks, California, London, 238.
- Morgan JF (2007). p Value fetishism and use of the Bonferroni adjustment. *Evid. Based Mental Health* 10: 34-35.
- Parzen E (1960). *Modern Probability Theory and its Applications*. xv. A Wiley Publication in Mathematical Statistics. Wiley, New York, 464.