HIGH RECOMBINATION AND MUTATION RATES IN MOUSE HEPATITIS VIRUS SUGGEST THAT CORONAVIRUSES MAY BE POTENTIALLY IMPORTANT EMERGING VIRUSES

Ralph S. Baric, Kaisong Fu, Wan Chen, and Boyd Yount

Department of Epidemiology
University of North Carolina
Chapel Hill, North Carolina 27599-7400

INTRODUCTION

Coronaviruses are common respiratory and gastrointestinal pathogens of mammals and birds. Not only do they cause about 15-20% of the common colds in humans, they are also occasionally associated with infections of the lower respiratory tract and central nervous system[1]. The prototype, mouse hepatitis virus (MHV), contains a 32 kb genomic RNA which encodes two large orfs at the 5′ end, designated orf 1a and orf 1b. Orf 1b contains highly conserved polymerase, helicase and metal binding motifs typical of viral RNA polymerases while orf 1a contains membrane and cysteine rich domains, and serine-and poliovirus 3c-like protease motifs[1]. The large size of the genome coupled with it's unique replication strategy and high recombination frequencies during mixed infection predict a considerable capacity to evolve[1,2,3].

The majority of emerging RNA viruses are probably zoonotic pathogens that bridge the species barrier and spread into the human population. The probable emergence of HIV, Hantaan, and influenza viruses from zoonotic hosts suggests that this is a natural, almost predictable phenomenon, yet we know little about the molecular mechanisms mediating virus spread between species[3]. To address this question in coronaviruses, we calculated MHV polymerase error rates and RNA recombination frequencies. Further MHV's capacity to bridge the species barrier was confirmed by isolating variants which grow efficiently on nonpermissive baby hamster kidney cells in vitro.

Corona- and Related Viruses, Edited by P. J. Talbot and G. A. Levy
Plenum Press, New York, 1995

MATERIAL AND METHODS

Temperature sensitive (*ts*) mutants from the group F RNA$^+$ mutants (NC6, NC16), group E (LA18) and group C (NC3) RNA$^-$ mutants were used in this study[2,3]. MHV-A59/JHM was adapted to baby hamster kidney cells (BHK) by serial passage in progressively increasing ratio's of BHK/DBT mixed cultures.

RESULTS

Sequence Analysis of TS and Revertent Viruses

Identification of the mutant allele in *ts* virus in MHV is complicated by the large size of the viral genome including the ~22 kb polymerase region encoding at least 5 or more genetic functions[2]. To simplify this problem, we focused on the RNA$^+$ mutants which mapped in the S gene of MHV (NC6, NC16) and the RNA$^-$ mutants which mapped in the n-(NC3) or c-termini (LA18) of orf 1b[2,3]. Since infectious vectors are not available, revertent viruses were isolated to assist in the identification of the mutant allele. Revertent viruses had similar titers and were of the RNA$^+$ phenotype when assayed at both permissive and restrictive temperatures (Table 1).

Overlapping primer pairs were developed to clone 1.0-1.5 kb stretches of viral RNA in the MHV orf 1b and S genes using PCR[5]. To obviate PCR-induced mutations and circumvent high RNA polymerase error rates[4], the PCR product was reamplified by asymmetric PCR, and sequenced directly. For the RNA$^+$ mutants (NC6, NC16), the entire S glycoprotein gene was sequenced in *ts* and revertent viruses. For the RNA$^-$ mutants (LA18, NC3), the entire orf 1b region and the c-terminal 1-2 kb of orf 1a was sequenced.

NC6 contained a single nucleotide substitution at the 5′ end of the S glycoprotein gene which resulted in an A to G transition at position 630 (Thr to Ala; amino acid 207). NC16 contained a single nucleotide change involving a C to G transversion at nt 2511 (Asp to Glu) in the S2 domain of the S glycoprotein gene. In revertents of NC16, the mutation had reverted back to wildtype sequence (Table 2).

Among the group E RNA$^-$ mutants, LA18 contained a single nucleotide substitution at nt 7100 in orf 1b. The G to A alteration led to a Arg to Lys change at amino acid 2286. In ten LA18 revertants, the mutation had reverted to the wildtype sequence. The group C mutant NC3 contained two alterations at nt 171 and 172; located upstream from the ribosomal

Table 1. *Ts* and revertant virus titers

	Titer at Temperature (°C)		
Ts mutant	32	39.5	39.5/32
LA18	6.1×10^7	2.4×10^3	4.0×10^{-5}
LA18R1	2.3×10^7	9.5×10^7	4.0×10^{-1}
NC3	2.4×10^8	5.5×10^4	2.3×10^{-4}
NC3R1	2.0×10^8	2.0×10^8	1.0×10^0
NC6	2.8×10^7	3.5×10^3	1.2×10^{-4}
NC6R1	2.3×10^7	2.4×10^7	1.0×10^0
NC16	8.5×10^7	2.3×10^4	2.7×10^{-4}

Table 2. Nucleotide locations of the MHV-A59 *ts* alleles

TS mutant	Location in genome	nt site	nt change		AA change	
			wt[a]	ts[b]/Rev[c]	wt	ts(Rev)
Group F						
NC6	S gene	620	A	G	Tyr	Cys
NC6R1		—	—	A	—	Tyr
NC16		2502	C	G	Asp	Glu
NC16R1		—	—	C	—	Asp
Group C						
NC3	ORF1a	171,172	GC	CG	Gly	Ala
NC3R1		—	—	GC	—	Gly
Group E						
LA18	ORF1b	7100	G	A	Arg	Lys
LA18R1		—	—	G	—	Arg

[a]wildtype.
[b]temperature sensitive mutant.
[c]revertant.

frameshifting site in orf 1a. The alteration (GC to CG) resulted in a Gly to Ala at amino acid 4438. In revertant virus, reversion to wildtype sequence had occurred.

Additional nucleotide changes were also found in *ts* and revertant viruses which were not tightly linked to the *ts* phenotype (data not shown).

Establishing Precise Recombination Rates in the MHV A59 Genome

All mutants were crossed 3-5 X to the other mutants used in the study and standardized to the standard cross (Table 3)[2,3]. To establish recombination frequencies between highly defined alleles, the nucleotide distances between individual mutants was divided by the recombination frequencies between each particular cross. Within the S gene (NC6 x NC16), the average recombination frequency was 1% RF/629±331 bp as compared to 1%/8979±1191 bp in the polymerase gene (LA18 x NC3). Between the 5′ end of the S glycoprotein gene (NC6) and the 3′ end of the polymerase gene (LA18), the average recombination frequency was about 1%/1054±139 bp, indicating that progressively increasing recombination frequencies were evident from the 5′ to 3′ end of the genome[3].

RNA Polymerase Error Rates during MHV Infection

MHV RNA polymerase error rates were calculated from the average reversion frequencies of several *ts* mutants and approached $1.32 \pm 0.89 \times 10^{-4}$ or about 2.4 mutations (range 1.5-7.2) per genome round of replication.

Table 3. Recombination frequencies between *ts* mutants[a]

Mutant	NC3	LA18	NC6	NC16
NC3	—	0.84 ± 0.21	5.50 ± 2.10	ND
LA18		—	0.90 ± 0.20	4.6 ± 0.60
NC6			—	2.8 ± 0.90
NC16				—

[a]mean ± standard deviation. ND, not done.

Table 4. Characterization of MHV BHK-adapted virus isolates

	Virus titers[a]	
	DBT	BHK
Virus stocks		
MHV-A59	7.3×10^8	0
MHV-H1	3.5×10^7	3.0×10^7
MHV-H2	1.0×10^8	7.5×10^7
Virus Growth[b]		
MHV-A59 (0h)	2.6×10^4	4.8×10^4
MHV-A59 (24h)	7.0×10^7	1.7×10^3
MHV-H2 (0h)	1.0×10^4	2.5×10^4
MHV-H2 (24h)	1.0×10^6	1.1×10^8

[a]virus stocks titered on DBT or BHK cells.
[b]cultures of DBT or BHK cells were infected with virus and titer on DBT cells.

MHV Evolution in Vitro

The high rate of recombination coupled with a high RNA polymerase error rate suggested that MHV has considerable potential to evolve rapidly in vitro. Since most emerging viruses are thought to be zoonotic pathogens that bridge the species barrier[4], we

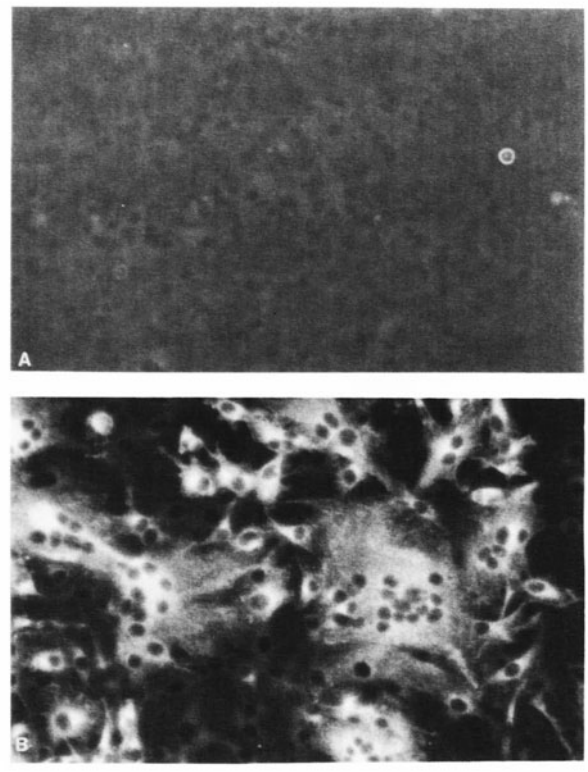

Figure 1. Expression of MHV-A59 (Panel A) and MHV-A59H2 (Panel B) viral antigen in BHK cells.

examined the capacity of MHV to bridge the species barrier and grow in nonpermissive BHK cells. Cultures of mixed BHK/DBT cells were infected with MHV and the progeny serially passaged in progressively increasing ratios of BHK to DBT cells (1 to 10, to 5 to 1). Following additional selection, two MHV variants (MHV-H1, MHV-H2) were isolated which replicated, expressed viral antigen, fused BHK cells, and plaque in BHK and DBT cells efficiently (Figure 1) (Table 4). Under identical conditions, MHV-A59 wild type virus did not replicate or express significant quantities of viral protein in BHK cells.

DISCUSSION

Ts and revertant viruses were sequenced to identify specific mutations which correlated with the *ts* phenotype, map the location of the MHV complementation groups, determine RNA polymerase error rates and precise RNA recombination frequencies throughout the genome. While this approach suffers from the practical limits of sequencing the entire 32 kb genome, each mutation was uniquely present in *ts* but not in wildtype or revertent viruses, and mapped to a domain as predicted by standard genetic recombination mapping techniques[2]. Although more *ts* mutants must be sequenced, these data suggest that the group F RNA^+ mutants map within the S glycoprotein gene of MHV while the group E RNA^- mutants, which are defective in positive strand synthesis after temperature shift, reside near the c-terminus of orf 1b. The group C mutants, which are defective in negative but not positive strand RNA synthesis, probably map at the orf 1a/orf 1b junction.

It is well documented that viral RNA-dependent RNA polymerases do not contain proofreading activities and have high mutation rates approaching 10^{-3} to 10^{-5} per round of replication[4]. The resulting mixed virus populations, quasi-species, is probably an important mechanism of RNA virus evolution and spread between species. From the reversion frequencies of highly plaque purified *ts* mutants, we determined the average MHV RNA polymerase error rate to approach $1.32 \pm 0.89 \times 10^{-4}$, or about 2.4 mutations per genome round of replication (range, 1.5-7.4). High mutation rates in a genome that is 3-4 X larger than most other positive-stranded RNA viruses suggests that coronaviruses probably exist in large quasispecies populations, providing vast reservoirs of natural virus variants.

Previous studies in our laboratory and others have suggested that RNA recombination occurs at high frequency and varies in different portions of the MHV genome[1,2,3]. The physical map locations of several MHV *ts* mutants provides accurate estimates of the homologous RNA recombination frequencies in the MHV genome which approach 1%/8979±1191 bp or about 4.9% throughout the 22 kb polymerase region (assuming reciprocal crosses). Between the polymerase and S glycoprotein genes, the recombination frequency is significantly increased, and approaches 1%/1054±139 bp (~3.8% in the 2.0 kb p30/HE genes). In the 4.0 kb S glycoprotein gene, RNA recombination rates approach 1%/629±331 bp (~12.7%). If recombination frequencies at the 3′ end of the genome approaches or exceeds rates measured in the S gene, the overall recombination rate in MHV probably exceeds 25%. Currently, the most likely model to explain the high, progressively 5′ to 3′ increased recombination rate in MHV is from: (1) the large size of the genome, (2) discontinuous transcription, and (3) the presence of transcriptionally active full and subgenomic length positive- and negative-stranded RNAs which increase the amount of template for strand switching[2,3].

The high rate of coronavirus RNA recombination coupled with high RNA polymerase error rates provides coronaviruses with a natural mechanism for rapid antigenic variation and evolution, especially within the highly immunogenic structural genes. It is clear that MHV can rapidly alter its species specificity and infect rats and primates; the resulting virus variants are associated with demyelinating diseases in these alternative species[5]. Newly

recognized animal coronaviruses include porcine epidemic diarrhea coronaviruses which probably evolved by mutation and recombination from the human coronavirus 229E[6]. We have also isolated MHV variants which efficiently replicate not only in BHK cells, but to a lesser extent, have retained their natural ability to replicate in mouse cells in vitro. These data further demonstrate the capacity of coronaviruses to evolve rapidly and bridge the species barrier in vitro. Using these variants, we can map determinants which mediate species specificity, map MHV virulence markers in the hamster and/or mouse, determine if the BHK CE homologue or other cellular proteins acts as a receptor in alternative hosts, and determine evolution rates in structural and nonstrucural genes. The data presented in this paper demonstrates that coronaviruses can evolve rapidly and transverse the species barrier in vitro.

ACKNOWLEDGMENTS

This research was supported by grants from the American Heart Association (AHA 90-1112), the NIH (AI-23946) and was performed during the tenure of an Established Investigator Award from the American Heart Association (RSB) (AHA 89-0193).

REFERENCES

1. Lai, M.M.C. 1990. Coronavirus: organization, replication and expression of genome. Annu. Rev. Microbiol. 44, 303-333.
2. Baric, R.S. Fu, K.S., Schaad, M.C. and Stohlman, S.A. 1990. Establishing a genetic recombination map for MHV-A59 complementation groups. Virology **177**: 646-656.
3. Fu, K. and Baric, R.S. 1992. Evidence for variable rates of recombination in the MHV genome. Virology **189**:88-102.
4. Morse, S.S. 1994. The viruses of the future? Emerging viruses and evolution. In: Morse, S.S. (eds), The Evolutionary Biology of Viruses, Raven Press, Ltd, New York.
5. Murray, R.S., et al., 1992. Coronavirus infects and causes demyelination in the primate central nervous system. Virology 188: 274-284.
6. Brigden, A., M. Duarte, K. Tobber, H. Laude, and M. Ackermann. 1993. The nucleocapsid protein gene of the porcine epidemic diarrhoea virus confirms that this virus is a coronavirus related to HCV 229E and TGEV. J. Gen. Virol. 74:1795-1804.