Analysis of the genomic homologous recombination in *Theilovirus* based on complete genomes

Guangming Sun1,2†, Xiaodan Zhang2†, Maoli Yi3, Shihe Shao2 and Wen Zhang2

**Abstract**

At present, *Theilovirus* is considered to comprise four distinct serotypes, including Theiler’s murine encephalomyelitis virus, Vilyuisk human encephalomyelitis virus, Thera virus, and Saffold virus. So far, there is no systematical study that investigated the genomic recombination of *Theilovirus*. The present study performed the phylogenetic and recombination analysis of *Theilovirus* over the complete genomes. Seven potentially significant recombination events were identified. However, according to the strains information and references related to the recombinants and their parental strains, four of the recombination events might happen non-naturally. These results will provide valuable hints for future research on evolution and antigenic variability of *Theilovirus*.

**Keywords:** *Theilovirus*, Recombination, Phylogenetic analysis

**Introduction**

*Encephalomyocarditis virus* (EMCV) and *Theilovirus* are two distinct species in the Cardiovirus genus of the family Picornaviridae [1]. The EMCVs comprise a single serotype and have a wide host range, while the *Theilovirus* species, probably includes four serotypes: Theiler’s murine encephalomyelitis virus (TMEV), Vilyuisk human encephalomyelitis virus (VHEV), Thera virus (TRV; isolated from rats) and Saffold virus (SAFV; isolated from humans). TMEVs were originally isolated from mice and later from rats [2]. Serological studies indicated that the feral house mouse *Mus musculus* is the natural host for TMEV [3]. VHEV was isolated by the inoculation of mice with nasopharyngeal secretions, serum samples, feces, cerebrospinal fluid (CSF) specimens, and brain specimens from the Yakut-Evenk population, indigenous rural people in Siberia that had a chronic form of encephalitis [4]. TRV was isolated from sentinel rats housed with TMEV-seropositive rats in Japan [5]. This virus has not yet been associated with disease in rats but has raised the possibility of additional clades of undiscovered theiloviruses.

SAFVs, new theiloviruses, were first isolated in California from a fecal sample from an 8-month-old infant with fever of undetermined origin [6] and then from a nasopharyngeal sample collected from a 23-month-old child in Canada in 2006 [7].

For picornaviruses, recombination is a common mechanism of evolution and antigenic variability. Although a recent report suggested that recombination happened in Cardiovirus genus [8], no study has systematically investigated the recombination among *Theilovirus* strains. In the present study, therefore, we systematically analyzed the available complete *Theilovirus* genome sequences in GenBank to elucidate the recombination among these viruses.

**Methods**

**Sequences**

The study sequences comprised all the 23 available complete genome sequences of *Theilovirus* from GenBank dated January 2011. Sequences were firstly screened to exclude patented and artificial mutants, and then aligned in the ClustalW program [9]. The alignment was manually adjusted for the correct reading frame. Sequences showing less than 1% divergence from each other were considered as the same. The strain information of the remaining 21 *Theilovirus* genomes...
were shown in Table 1. Because there was no complete genome of VFHV in GenBank before our analysis, this virus was not analyzed in the present study.

Phylogenetic analysis
Before phylogenetic analysis, multiple-alignment was performed in the ClustalW program. Phylogenetic trees were constructed using the neighbor-joining method and evaluated using the interior branch test method with Mega 4 software [10]. Percent bootstrap support was indicated at each node. GenBank accession no. was indicated at each branch.

Recombination Detection
The remaining 21 Theilovirus genomes were re-aligned in the ClustalW program. Detection of potential recombinant sequences, identification of potential parental sequences, and localization of possible recombination break points were determined using the Recombination Detection Program (RDP) [11], GENECONV [12], BOOTSCAN [13], MaxChi [14], CHIMAERA [15], and SISCAN [16] methods embedded in RDP3 [17]. A multiple-comparison-corrected P-value cutoff of 0.05 was used throughout.

Results and Discussion
Based on the 21 complete Theilovirus genomes, a phylogenetic tree was constructed (Figure 1). The taxonomy of these Theilovirus showed in the phylogenetic tree was consistent with the strain information from the original sources. From the phylogenetic tree, we can see that Theilovirus were divided into two major different genetic groups. Among the two major groups, SAFV formed a single group, while TMEV and TRV closely clustered, forming the other group. Sequence alignment indicated that TMEV strains shared 71.2%-75.3% and 67.4%-70.1% sequence identities with TRV and SAFV strains, respectively. While TRV strains showed 72.2%-74.8% sequence homologies to SAFV strain.

Seven potentially significant recombination events were detected with a high degree of confidence (p value $\leq 1.3 \times 10^{-4}$) judged by the above-mentioned six recombination detection methods. Figure 2 indicated the 7 recombination events, where we can see that event1 included three recombinants which had the same parental strains while the other six recombination events contained six recombinants, respectively.
Figure 3 showed the identification result of recombination event 1, which occurred between the lineage represented by a Germany SAFV strain [GenBank: EU681177] [18] as the minor parent and a USA SAFV strain [GenBank:EF165067] [6] as the major parent. This recombination event led to three recombinant SAFV strains [GenBank:EU376394, EMBL:AM922293, [GenBank:GU595289 ]\[7,19,20\]. In this recombination event,
the two parental strains were isolated in different countries, and the three daughter recombinants were distributed in different countries, which might hint that this recombination event happened long time ago and the recombinants were prevalent worldwide.

Recombination event2 identified the recombination occurred between two SAFV strains [GenBank: GU595289, GenBank:EU681179], leading to the other recombinant SAFV strain [GenBank:EU681176] (Additional File 1, Part A). However, in this recombination event, one of the parental strain [GenBank:EU681179] and the daughter strain were sequenced in the same lab [18], therefore, whether this recombination event occurred naturally or not should be verified by future studies. Additional File 1, Part B and C indicated the recombination event3 and event4, respectively, and three SAFV strains [GenBank: FJ463615, GenBank:FJ463616, GenBank:FJ463617] involved in the two recombination events were all sequenced in the same lab [21], therefore, it should be cared whether these two recombination events non-naturally occurred by sequencing error and/or contamination. The recombination event5 (Additional File 1, Part D) also contained two strains [GenBank: EU681179, GenBank: EU681178] which were isolated in the same lab [18], therefore, whether this recombination event non-naturally occurred by sequencing error and/or contamination should be elucidated by further study.

Figure 4 indicated the recombination event6 that occurred between a two TMEV strains, Yale strain [GenBank:EU723238] and DA strain [GenBank: M20301] [22], which led to the recombinant TMEV strain BeAn [GenBank:M16020] which was isolated from mouse in 1987, and these three virus strains were all isolated from mouse in USA [1,22]. Figure 5 revealed the putative recombinant TMEV strain (GenBank:M20301), however, the accurate parental strains has not been detected in the present study, which may due to the limited numbers of Theilovirus sequence available at present, therefore, further study should be performed to identify the accurate parental strains with the increasing number of Theilovirus genome sequences.

Recombination is a relatively common phenomenon in RNA viruses and understanding recombination will be helpful in unravelling the evolution of pathogens and drug resistance [23-25]. In the present study, we performed phylogenetic and recombination analyses over the full genome of Theilovirus available in GenBank nowadays. Seven potentially significant recombination events were detected. However, four of the
recombination events might happen non-naturally in the lab, which should be taken into notice in the future evolutionary analysis of *Theilovirus*. The other three recombination events were further analyzed using other algorithms in RDP software bag and some of them were confirmed by phylogenetic analysis. The recombination phenomena of *Theilovirus* will also be noted in the further research because this will be one pattern of virulence factor variation in *Theilovirus*.

**Additional material**

Additional file 1: **BOOTSCAN evidence for the recombination event** 2, 3, 4, and 5. Analysis was based on pairwise distance, modeled with a window size 200, step size 20, and 100 Bootstrap replicates.

**References**

1. Liang Z, Kumar AS, Jones MS, Knowles NJ, Lipton HL. Phylogenetic analysis of the species *Theilovirus*: emerging murine and human pathogens. *J Virol* 2008, 82:11545-11554.
2. Hermel E, Huxsoll DL, Wammer AR Jr. Comparison of MHV virus with mouse encephalomyelitis viruses. *Lab Anim Sci* 1974, 24:523-529.
3. Desclèques JP, Mihok S. Serologic study on the prevalence of murine viruses in a population of wild meadow voles (*Microtus pennsylvanicus*). *J Wildl Dis* 1986, 22:314-319.
4. Goldfus LG, Gajdusek DC. Villiush encephalomyelitis in the lakut people of Siberia. *Brain* 1992, 115:961-978.
5. Ohawa K, Watanabe Y, Miyata H, Sato H. Genetic analysis of a Theiller-like virus isolated from rats. *Comp Med* 2003, 53:191-196.
6. Jones MS, Lukashov V, Gnan RD, Schnurr DP. Discovery of a novel human picornavirus in a stool sample from a pediatric patient presenting with fever of unknown origin. *J Clin Microbiol* 2007, 45:2144-2150.
7. Abed Y, Boivin G. *New Saffold cardioviruses in 3 children, Canada*. *Emerg Infect Dis* 2008, 14:834-836.
8. Drexler JF, Baumgarte S, Luna UK, Stocker A, Almeida PS, Ribeiro TC, Petersen N, Herzog P, Pedroso C, Brites C, Ribeiro Hda C Jr, Gryn A, Drosten C, Lukashev A. Genomic features and evolutionary constraints in Saffold-like cardioviruses. *J Gen Virol* 2010, 91:1418-1427.
9. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994, 22:4673-4680.
10. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007, 24:1596-1599.
11. Martin D, Rubycki E. RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 2000, 16:562-563.
12. Paddam M, Sawyer S, Fauquet CM. Possible emergence of new geminiviruses by frequent recombination. *Virology* 1999, 265:218-225.
13. Martin DP, Posada D, Crandall KA, Williamson C. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Res Hum Retrov* 2005, 21:98-102.
14. Smith JM. Analyzing the mosaic structure of genes. *J Mol Evol* 1992, 34:126-9.
15. Posada D, Crandall KA. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc Natl Acad Sci USA* 2001, 98:13737-13742.
16. Gibbs MJ, Armstrong JS, Gibbs AJ. Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* 2000, 16:573-582.
17. Martin DP, Williamson C, Posada D. *RDP2*: recombination detection and analysis from sequence alignments. *Bioinformatics* 2005, 21:260-262.
18. Drexler JF, Luna UK, Stocker A, Almeida PS, Ribeiro TC, Petersen N, Herzog P, Pedroso C, Huppert H, Ribeiro Hda C Jr, Baumgarte S, Drosten C. Circulation of 3 lineages of a novel Saffold cardiovirus in humans. *Emerg Infect Dis* 2008, 14:1398-13405.
19. Chiu CY, Greninger AL, Chen EC, Haggerty TD, Parsonnet J, Delwart E, Derisi JL, Gamem D. Cultivation and serological characterization of a human Theiller-like cardiovirus associated with diarrheal disease. *J Virol* 2010, 84:4407-4414.
20. Chiu CY, Greninger AL, Kanada K, Kwock T, Fischer FK, Runkel C, Louie JK, Gaiser CA, Yagi S, Schnurr DP, Haggerty TD, Parsonnet J, Gamem D, Derisi JL. Identification of cardioviruses related to Theiller's murine encephalomyelitis virus in human infections. *Proc Natl Acad Sci USA* 2008, 105:14124-14129.
21. Blinkova O, Kapoor A, Victoria J, Jones M, Wolfe N, Naeem A, Shaukat S, Shafir S, Alam MW, Angez M, Zaidi S, Delwart EL. Cardioviruses are genetically diverse and cause common enteric infections in South Asian children. *J Virol* 2009, 83:4631-4641.
22. Ohana Y, Stein S, Fu JL, Stillman L, Klamann L, Roos RP. Molecular cloning and sequence determination of DA strain of Theiller's murine encephalomyelitis viruses. *Virology* 1988, 164:245-253.
23. Wang H, Zhang W, Ni B, Shen H, Song Y, Wang X, Shao S, Hua X, Cui L. Recombination analysis reveals a double recombination event in hepatitis E virus. *J Virol* 2010, 71:29.
24. Pickett BE, Lefkowitz EJ: Recombination in West Nile Virus: minimal contribution to genomic diversity. Virol J 2009, 6:165.

25. Moreno P, Alvarez M, López L, Moratorio G, Casane D, Castells M, Castro S, Cristina J, Colina R: Evidence of recombination in Hepatitis C Virus populations infecting a hemophiliac patient. Virol J 6:203.

doi:10.1186/1743-422X-8-439

Cite this article as: Sun et al.: Analysis of the genomic homologous recombination in Theilovirus based on complete genomes. Virology Journal 2011 8:439.

Submit your next manuscript to BioMed Central and take full advantage of:

• Convenient online submission
• Thorough peer review
• No space constraints or color figure charges
• Immediate publication on acceptance
• Inclusion in PubMed, CAS, Scopus and Google Scholar
• Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit