VIROLOGY JOURNAL

RESEARCH Open Access

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

Guangming Sun[1,2*†], Xiaodan Zhang[2†], Maoli Yi[3], Shihe Shao[2] and Wen Zhang[2]

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

* Correspondence: xhxsgm@yahoo.com
† Contributed equally
[1]The Fourth Affiliated Hospital of Jiangsu University, 20 Zhengdong Road, Zhenjiang, Jiangsu 212001, China
Full list of author information is available at the end of the article

Table 1 The 21 *Theilovirus* strains used in phylogentic and recombination analysis in the present study

GenBank No.	Strain Name	Source	Country	Virus
M16020	BeAn	Mouse	USA	TMEV
M20301	DA	Mouse	USA	TMEV
X56019	GDVII	Mouse	UK	TMEV
EU718733	Vie415HRT	Mouse	USA	TMEV
EU723238	Yale	Mouse	USA	TMEV
EU718732	TOB15	Mouse	USA	TMEV
AB090161	NGS910	Rat	Japan	TRV
EU542581	TRV-1	Rat	USA	TRV
EU815052	RTV1	Rat	USA	TRV
EU681178	D/VI2273/2004	Human	Germany	SAFV
EU681179	D/VI2223/2004	Human	Germany	SAFV
FJ463616	Pak5152	Human	Pakistan	SAFV
FJ463615	Pak5003	Human	Pakistan	SAFV
FJ463617	Pak6572	Human	Pakistan	SAFV
EF165067	NA	Human	USA	SAFV
EU681177	BR/118/2006	Human	Germany	SAFV
EU681176	D/VI2229/2004	Human	Germany	SAFV
GU595289	HTCV-UC6	Human	USA	SAFV
EU376394	NA	Human	USA	SAFV
AM922293	Can112051-06	Human	Canada	SAFV
FM207487	Nijmegen2007	Human	Netherlands	SAFV

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

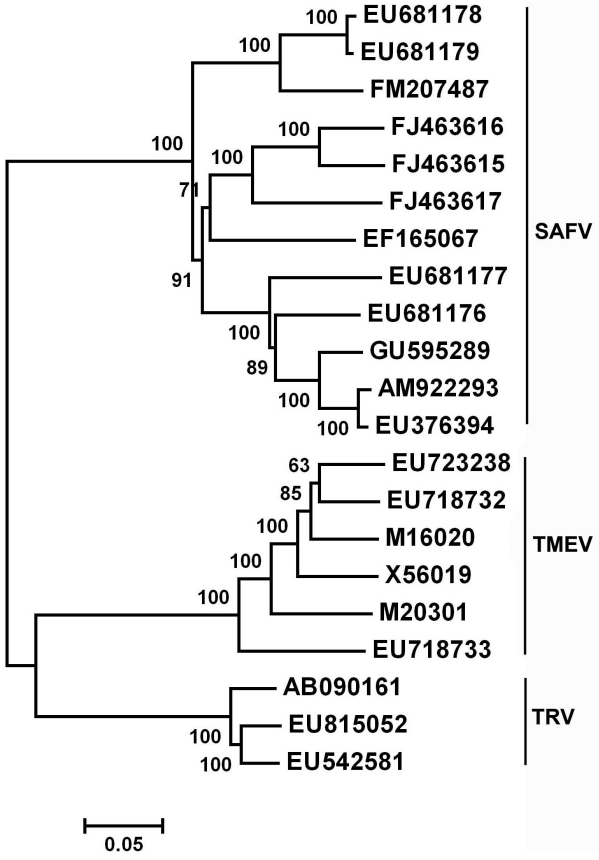

Figure 1 Phylogenetic tree for the 21 complete *Theilovirus* genomes. Phylogenetic analysis were performed using the neighbor-joining method and evaluated using the interior branch test method with Mega 4 software. Values for various branches are percentages of the tree obtained from 1000 resamplings of the data. Percent bootstrap supports are indicated at nodes.

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 genetical 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.

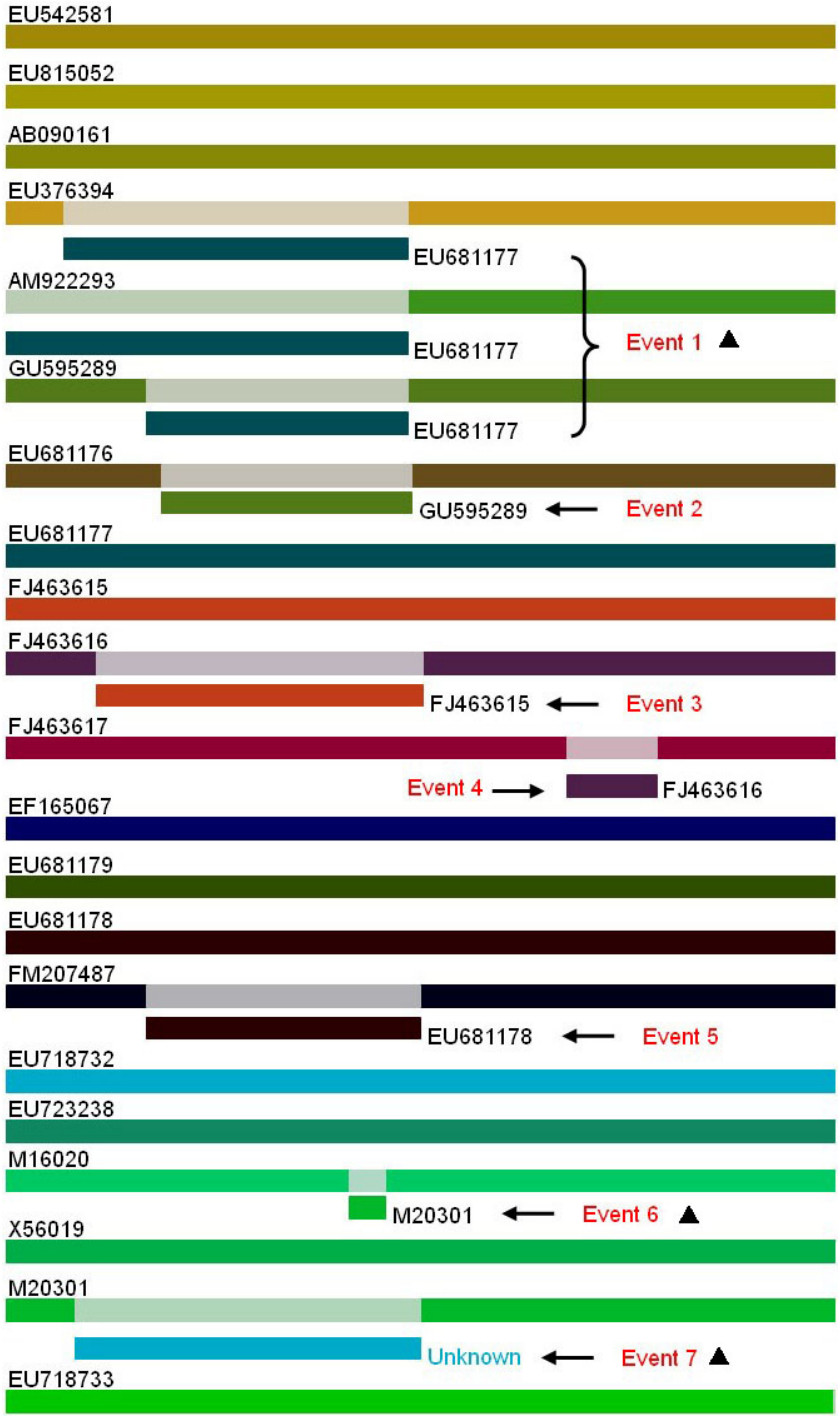

Figure 2 Identification of the 7 recombination events. The recombination events were indicated in red word "event"; GenBank No. of each strain was indicated at the left end; the minor parental strain of each recombinant was shown at the recombination region. The solid triangles indicated the naturally occurred recombination events.

Figure 3 showed the identification result of recombination event1, 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,

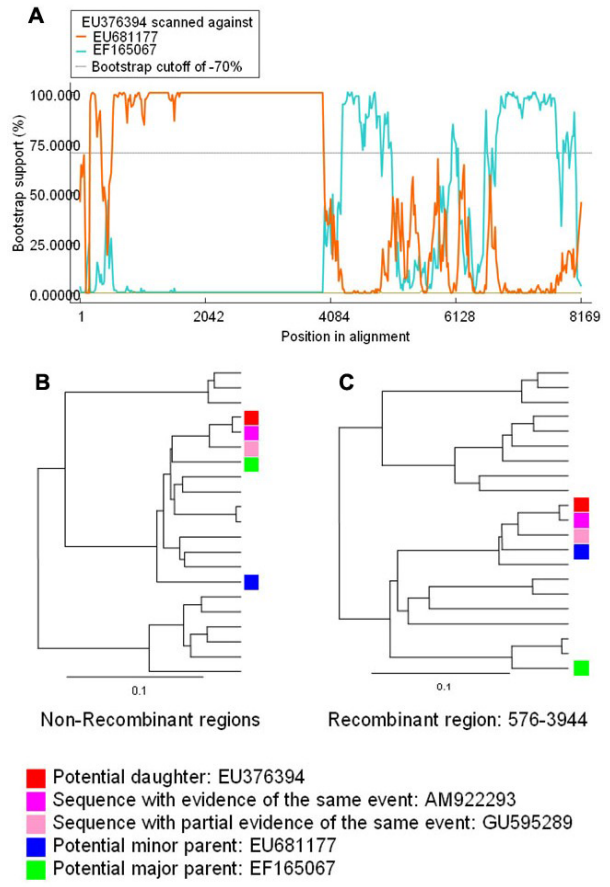

Figure 3 Identification of recombination between EU681177 and EF165067. (A) BOOTSCAN evidence for the recombination origin on the basis of pairwise distance, modeled with a window size 200, step size 20, and 100 Bootstrap replicates; (B) Neighbor joining tree (2,000 replicates, Kimura 2-parameter distance) constructed using the non-recombinant region (Position 1-575 + 3945-end); (C) Neighbor joining tree (2,000 replicates, Kimura 2-parameter distance) constructed using the recombinant region (Position 576-3944).

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 [19], 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

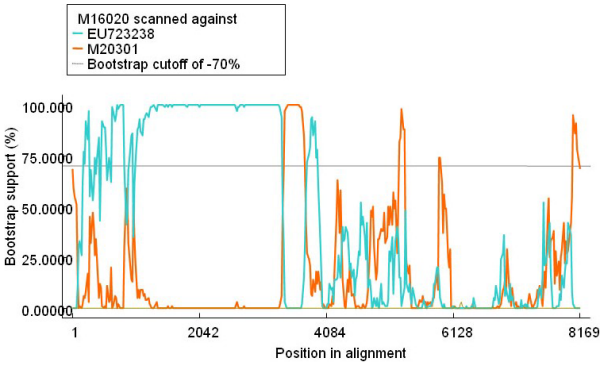

Figure 4 BOOTSCAN evidence for recombination between two TMEV strains, which led to a recombinant TMEV strain. Analysis were based on the basis of pairwise distance, modeled with a window size 200, step size 20, and 100 Bootstrap replicates.

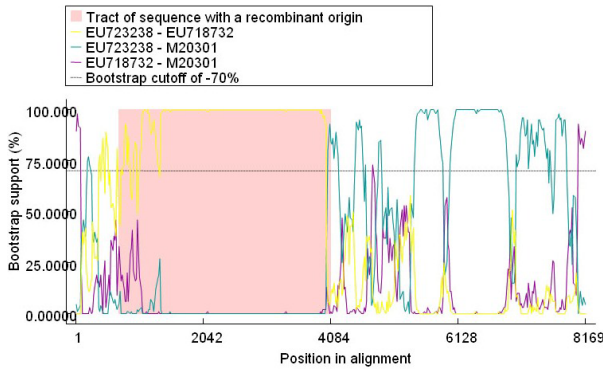

Figure 5 RDP screenshots displaying the possible recombinant (GenBank:M20301). The y-axis indicates the pairwise identity that refers to the average pairwise sequence identity within a 30nt sliding window moved one nucleotide at a time. The area outlined in gray demarcates the potential recombinant regions.

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.

Acknowledgements
This work was supported by Foundation for Society Development Schedule of Zhenjiang City under Grant No.2010041, and the Professional Research Foundation for Advanced Talents of Jiangsu University under Grant No.10JDG059.

Author details
[1]The Fourth Affiliated Hospital of Jiangsu University, 20 Zhengdong Road, Zhenjiang, Jiangsu 212001, China. [2]School of Medical Science and Laboratory Medicine, Jiangsu University, 301 Xuefu Road, Zhenjiang, Jiangsu 212013, China. [3]Yantai Yuhuangding Hospital, 20 Yudong Road, Yantai, Shandong 264000, China.

Authors' contributions
GS conceived the study. All authors performed recombination analysis, critically reviewed, and approved the final manuscript. GS wrote the paper. All authors read and approved the final manuscript

Competing interests
The authors declare that they have no competing interests.

Received: 20 August 2011 Accepted: 17 September 2011
Published: 17 September 2011

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. Hemelt IE, Huxsoll DL, Warner AR Jr: **Comparison of MHG virus with mouse encephalomyelitis viruses.** *Lab Anim Sci* 1974, **24**:523-529.
3. Descôteaux 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. Goldfarb LG, Gajdusek DC: **Viliuisk encephalomyelitis in the lakut people of Siberia.** *Brain* 1992, **115**:961-978.
5. Ohsawa K, Watanabe Y, Miyata H, Sato H: **Genetic analysis of a Theiler-like virus isolated from rats.** *Comp Med* 2003, **53**:191-196.
6. Jones MS, Lukashov VV, Ganac 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 LK, Stöcker A, Almeida PS, Ribeiro TC, Petersen N, Herzog P, Pedroso C, Brites C, Ribeiro Hda C Jr, Gmyl A, Drosten C, Lukashev A: **Genomic features and evolutionary constraints in Saffold-like cardioviruses.** *J Gen Virol* 2010, **91**:1418-4127.
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, Rybicki E: **RDP: detection of recombination amongst aligned sequences.** *Bioinformatics* 2000, **16**:562-563.
12. Padidam 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 Retrovir* 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**:13757-13762.
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 LK, Stöcker A, Almeida PS, Ribeiro TC, Petersen N, Herzog P, Pedroso C, Huppertz HI, 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, Ganem D: **Cultivation and serological characterization of a human Theiler's-like cardiovirus associated with diarrheal disease.** *J Virol* 2010, **84**:4407-4414.
20. Chiu CY, Greninger AL, Kanada K, Kwok T, Fischer KF, Runckel C, Louie JK, Glaser CA, Yagi S, Schnurr DP, Haggerty TD, Parsonnet J, Ganem D, DeRisi JL: **Identification of cardioviruses related to Theiler'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, Sharif S, Alam MM, 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. Ohara Y, Stein S, Fu JL, Stillman L, Klaman L, Roos RP: **Molecular cloning and sequence determination of DA strain of Theiler's murine encephalomyelitis viruses.** *Virology* 1988, **164**:245-255.
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.** *Virol J* 2010, **7**:129.

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.