Virus Evolution, 2024, **10(1)**, veae054
DOI: https://doi.org/10.1093/ve/veae054
Advance Access Publication 20 July 2024
Research Article

Deciphering the evolutionary landscape of severe fever with thrombocytopenia syndrome virus across East Asia

Dongbin Park[1], Kwan Woo Kim[1], Young-Il Kim[1], Mark Anthony B. Casel[1,2], Hyunwoo Jang[1,2], Woohyun Kwon[1,2], Kanghee Kim[1,2], Se-Mi Kim[1], Monford Paul Abishek N[1], Eun-Ha Kim[3], Hobin Jang[1], Suhee Hwang[1], Seok-Min Yun[4], Joo-Yeon Lee[4], Hye Won Jeong[2], Su-Jin Park[5], Young Ki Choi[1,2,*]

[1]Center for Study of Emerging and Re-emerging Viruses, Korea Virus Research Institute, Institute for Basic Science (IBS), Daejeon, Republic of Korea
[2]College of Medicine and Medical Research Institute, Chungbuk National University, Cheongju, Republic of Korea
[3]Virus Research Resource Center, Korea Virus Research Institute, Institute for Basic Science (IBS), Daejeon, Republic of Korea
[4]Center for Emerging Virus Research, National Institute of Infectious Diseases, Korea National Institute of Health, Korea Disease Control and Prevention Agency, Cheongju, Republic of Korea
[5]Divison of Life Science, Research Institute of Molecular Alchemy (RIMA), Gyeongsang National University, Jinju, Republic of Korea

*Corresponding author. Center for Study of Emerging and Re-emerging Viruses, Korea Virus Research Institute, Institute for Basic Science (IBS), Daejeon 34126, Republic of Korea; College of Medicine and Medical Research Institute, Chungbuk National University, Cheongju 28644, Republic of Korea.
E-mail: choiki55@ibs.re.kr

Abstract

Severe fever with thrombocytopenia syndrome virus (SFTSV) poses a significant public health challenge in East Asia, necessitating a deeper understanding of its evolutionary dynamics to effectively manage its spread and pathogenicity. This study provides a comprehensive analysis of the genetic diversity, recombination patterns, and selection pressures across the SFTSV genome, utilizing an extensive dataset of 2041 sequences from various hosts and regions up to November 2023. Employing maximum likelihood and Bayesian evolutionary analysis by sampling trees (BEAST), we elucidated the phylogenetic relationships among nine distinct SFTSV genotypes (A, B1, B2, B3, B4, C, D, E, and F), revealing intricate patterns of viral evolution and genotype distribution across China, South Korea, and Japan. Furthermore, our analysis identified 34 potential reassortments, underscoring a dynamic genetic interplay among SFTSV strains. Genetic recombination was observed most frequently in the large segment and least in the small segment, with notable recombination hotspots characterized by stem-loop hairpin structures, indicative of a structural propensity for genetic recombination. Additionally, selection pressure analysis on critical viral genes indicated a predominant trend of negative selection, with specific sites within the RNA-dependent RNA polymerase and glycoprotein genes showing positive selection. These sites suggest evolutionary adaptations to host immune responses and environmental pressures. This study sheds light on the intricate evolutionary mechanisms shaping SFTSV, offering insights into its adaptive strategies and potential implications for vaccine development and therapeutic interventions.

Keywords: severe fever with thrombocytopenia syndrome virus; evolution; viral evolution; phylogenetic analysis; reassortment; recombination; selection.

Introduction

Severe fever with thrombocytopenia syndrome (SFTS) is an emerging infectious disease that poses significant public health challenges across East and Southeast Asia. Characterized by a wide range of symptoms from mild fever to severe hemorrhages and multiple organ failures, SFTS exhibits variable mortality rates between 5.1% and 18.7% influenced by geographic variations (Casel et al. 2021, Kim et al. 2023). Identified first in China in 2009 and later across other Asian territories, the spread and fatality of SFTS underscore the urgent need for comprehensive understanding and control measures (Yu et al. 2011, Tran et al. 2019). Transmission of SFTS primarily occurs through tick bites, notably by *Haemaphysalis longicornis*, with additional potential for human-to-human transmission via bodily fluids, presenting a complex challenge in managing disease spread (Gai et al. 2012, Zhuang et al. 2018).

The causative agent of SFTS, the SFTS virus (SFTSV), is a segmented RNA virus belonging to the genus *Bandavirus* within the family *Phenuiviridae* (Sasaya et al. 2023). Its genome is divided into large (L), medium (M), and small (S) segments, encoding critical viral proteins (Liu et al. 2023). Phylogenetic studies have identified eight distinct SFTSV genotypes, A through F, with geographical variances in dominance, indicating a diverse viral evolution pattern across regions (Fu et al. 2016; Yun et al. 2020).

Genetic reassortment and recombination within RNA viruses are crucial evolutionary mechanisms, allowing the emergence

of novel and potentially more virulent strains. Although segmented RNA viruses like SFTSV inherently facilitate reassortment, the incidence and impact of recombination within such viruses remain less understood, given the traditionally lower recombination rates observed in negative-strand RNA viruses (Perez-Losada et al. 2015, McDonald et al. 2016). These evolutionary processes contribute significantly to the viral genetic diversity, affecting the disease epidemiology and the development of vaccines and treatments.

This study aims to elucidate the complex evolutionary dynamics of SFTSV through detailed genetic and phylogenetic analyses of full-length viral genomes from infected patients and tick pools, alongside 2041 strains of complete open reading frame (ORF) sequences from the National Center for Biotechnology Information (NCBI) GenBank database. Our investigation reveals 119 recombination events across all three genomic segments and identifies 34 reassortment genotypes, indicating active viral evolution through these mechanisms. By charting the chronological sequence of reassortments and pinpointing genetic recombination hotspots, we highlight the ongoing adaptation and diversification of SFTSV, underscoring the critical need for continuous surveillance and advanced research to mitigate the public health impact of this emerging pathogen.

Materials and methods

SFTSV genome database compilation

A comprehensive dataset of complete SFTSV coding sequences, along with relevant metadata, was compiled from the NCBI GenBank database as of 23 November 2023. Sequences were analyzed and annotated using PfamScan and aligned with reference sequences (NC_018136, NC_018138, and NC_018137) to ensure accuracy (Madeira et al. 2022). To clarify the source information, our study excluded gene sequences if a sample did not provide collection dates and countries. A total of 1614 L, 1655 M, and 1798 S segment sequences were included for further analysis. Specific information on each sequence can be found in Supplementary Table S1.

Phylogenetic analysis

To elucidate the evolutionary relationships among SFTSV genotypes, representative strains embodying the recognized major genotypes were meticulously selected. Sequences from the L, M, and S genomic segments were clustered at specific identity thresholds—97.5% for L, 97% for M, and 97% for S—utilizing CD-HIT software (version 4.8.1) to facilitate this selection (Fu et al. 2012). For the representative strains, the RNA-dependent RNA polymerase (RdRp) genes from the L segment and the glycoprotein genes from the M segment were analyzed. Concurrently, the nucleoprotein (NP) and nonstructural (NS) genes of the S segment were concatenated for comprehensive evaluation. This preparatory step was succeeded by multiple sequence alignment using MAFFT software (version 7.505) (Rozewicki et al. 2019), paving the way for the inference of maximum likelihood (ML) phylogenetic trees with IQ-TREE (version 1.6.12) (Nguyen et al. 2015). The resultant phylogenetic trees were concisely summarized utilizing the MEGA11 software (Tamura et al. 2021).

To rigorously analyze SFTSV's molecular evolution, we generated phylogenetic trees using Bayesian Evolutionary Analysis by Sampling Trees (BEAST) software (version 1.10.4) that incorporate temporal dynamics (Suchard et al. 2018). We applied the general time reversible nucleotide substitution model augmented with invariable sites (I) and gamma-distributed rate variation (Γ). This was paired with a nonparametric Bayesian skygrid coalescent tree prior to handling flexible changes in population size among lineages effectively. The molecular clock model utilized an uncorrelated relaxed clock with a lognormal distribution. Our Bayesian phylogenetic analyses were conducted over Markov chain Monte Carlo chains run for 100 000 000 iterations, sampling every 50 000 steps. We assessed the convergence of these chains and the adequacy of the sample sizes using Tracer (version 1.7.2) (https://beast.community/tracer) to ensure reliable estimates. After discarding the first 10% of trees as burn-in, we used the remaining data to generate maximum clade credibility (MCC) trees using TreeAnnotator (version 1.10.4) (https://beast.community/treeannotator). These trees were visualized using FigTree (version 1.4.4) (http://tree.bio.ed.ac.uk/software/figtree/), facilitating detailed examination of phylogenetic relationships. Estimates of evolutionary rates and the times to the most recent common ancestor (tMRCAs) for each segment were reported as medians along with their 95% highest posterior density (HPD) intervals.

Reassortment and recombination analysis

In order to elucidate the dynamics of SFTSV genetic reassortment, a subset of 1513 strains, each encapsulating the complete tripartite genomic architecture (L, M, and S segments), was meticulously selected from an extensive dataset of 2041 strains. The analytical foundation for identifying potential reassortants was the ML phylogenetic tree. Strains exhibiting genotype discrepancies across the segmented genomes were flagged as candidates for reassortment, reflecting the virus's capacity for genetic exchange and evolution.

Further scrutiny was applied to uncover potential recombination events across the SFTSV genome spectrum. This investigation was conducted utilizing the Recombination Detection Program version 5 (RDP5) software, a comprehensive suite designed for the detection of viral recombination (Martin et al. 2021). The software employs a battery of algorithms—namely, RDP, GENECONV, Bootscan, Maxchi, Chimera, SiScan, and 3Seq—to rigorously analyze sequence data for recombination signals. To ensure robustness and reliability in the detection of recombination events, only those instances corroborated by at least three distinct algorithms, coupled with statistically significant *P*-value thresholds, were classified as *bona fide* recombinant events.

RNA secondary structure prediction

VaRNA (http://varna.lri.fr/) was used to predict the secondary structures of RNA (Darty et al. 2009). In brief, the Co-fold database was utilized to generate the dot parentheses of the input sequence for analysis (Proctor and Meyer 2013). Turner's nearest-neighbor model was followed to identify the structures using VaRNA, classifying them as nearest-neighbor loops like hairpin loops, internal loops, and bulge loops.

Selection pressure analysis

Selection pressure analysis for viral sequences was carried out on each segment of SFTSV using hypothesis testing using phylogenies (HyPhy) (version 2.5.24) (Kosakovsky Pond et al. 2020). The nonsynonymous (dN) versus synonymous (dS) nucleotide ratio (dN/dS) within the sequence was calculated using the single-likelihood ancestor counting (SLAC) implemented in HyPhy. In general, codons were classified as neutral ($dN/dS = 1$), undergoing positive selection ($dN/dS > 1$), or experiencing purifying selection ($dN/dS < 1$). Furthermore, sites exposed to positive selection pressure were also identified using the Fast Unbiased

Bayesian AppRoximation (FUBAR), mixed effects model of evolution (MEME), and fixed-effect likelihood (FEL) methods. In the SLAC, FEL, and MEME methods, only sites showing a statistically significant *P*-value (<.05) were considered under positive selection, while in the FUBAR method, only sites with a posterior probability of ≥.90 were deemed to be under positive pressure.

Results

Comprehensive genetic analysis of SFTSV strains across diverse hosts and geographies

To perform an in-depth analysis of the evolutionary relationships among SFTSV genotypes, we collected a dataset of 2041 SFTSV strains from the GenBank database, augmented with our recent sequencing efforts, covering various hosts until 23 November 2023. This collection comprises 1614 strains for the L segment, 1655 for the M segment, and 1798 for the S segment (Supplementary Table S1, Fig. 1a). A predominant portion of these strains (1862 out of 2003) was isolated from humans, followed by companion animals—primarily cats ($n = 56$) and dogs ($n = 15$) (Fig. 1b). Wild animals provided a smaller subset of the data, with strains from raccoons ($n = 4$), hedgehogs ($n = 3$), sheep ($n = 3$), and others. Although ticks are widely recognized as the primary vectors for SFTSV transmission, our dataset also includes an isolate from other arthropods. Specifically, it comprises 48 strains from tick hosts, primarily *H. longicornis* ($n = 38$), followed by *H. formosensis* ($n = 4$), and one strain isolated from a horsefly. This inclusion reflects the broader diversity of arthropods involved in SFTSV ecology. The temporal analysis of the virus sequencing data indicated the initial documentation of two strains in 2009, with a notable annual increase peaking in 2014 at 288 strains. From then until 2022, the annual strain report ranged from 25 to 258 (Fig. 1c). Despite year-round detection, a clear seasonality in SFTSV isolation was observed from March to October, reflecting a seasonal trend in virus prevalence. Geographically, the majority of SFTSV strains, accounting for 77.8% (1223 sequences), originated from China, with South Korea (15.8%, 249 sequences) and Japan (6.1%, 96 sequences) also contributing significantly (Fig. 1d). This finding is supported by the identification of partial SFTSV sequences in several countries across Southeast and South Asia (Win et al. 2020, Wu et al. 2021), highlighting the virus's extensive presence even in regions where complete genome data are lacking.

Genetic diversity and evolutionary insights of SFTSV across multiple hosts and regions

Among the 2041 SFTSV strains collected, 1513 possessed complete data across all three genomic segments (L, M, and S), forming a solid foundation for an in-depth phylogenetic analysis. In our extensive examination of the evolutionary dynamics among SFTSV genotypes, we meticulously selected representative strains for each of the major genotypes identified for these segments. Employing the CD-HIT software (version 4.8.1) and setting identity thresholds at 97.5% for L, 97% for M, and 97% for S, we effectively captured the wide range of genetic diversity within the SFTSV virus population. This strategic approach ensured that our analysis comprehensively represented the genomic variability of SFTSV. Our analysis unveiled a minimum nucleotide sequence similarity of 93.4% and an amino acid sequence similarity of 94.8% across the SFTSV genomes. This indicates a relatively modest genetic diversity among the strains (Wang et al. 2023). This trend of similarity was consistent across all segments: L segments displayed nucleotide and amino acid similarities ranging from 99.7% to 95.5% and 99.9% to 98.4%, respectively. M segments showed nucleotide similarities from 99.5% to 93.4% and amino

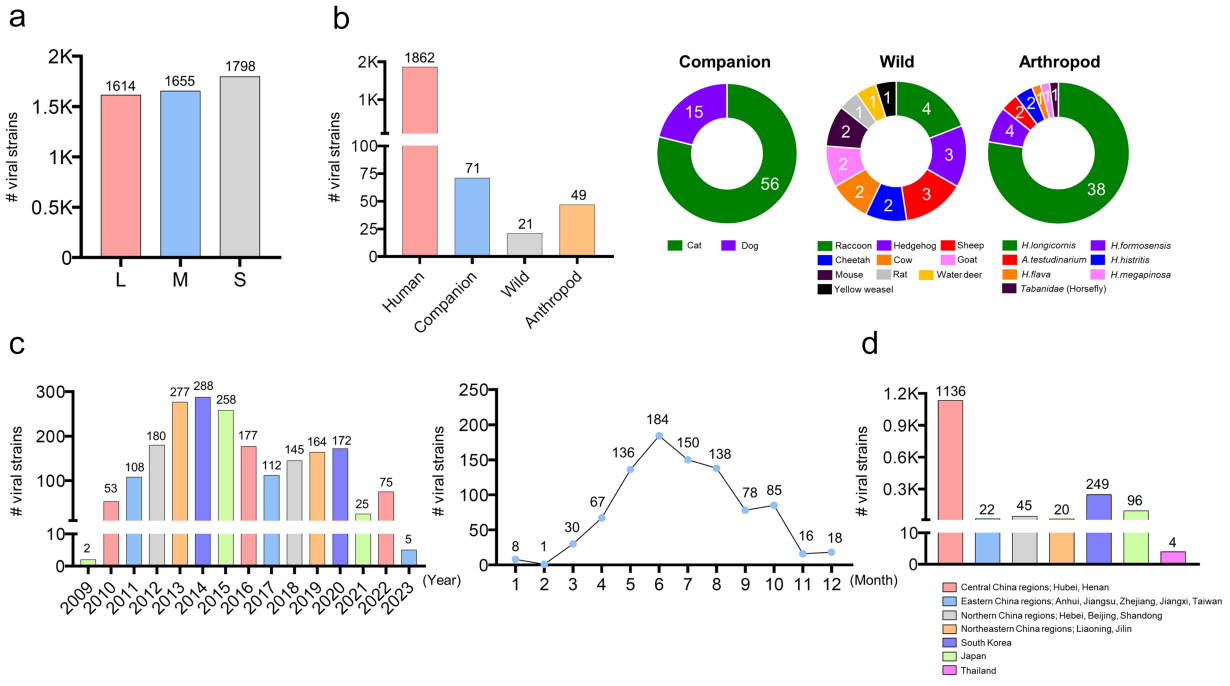

Figure 1. Information on SFTSV strains analyzed in this study.

(a) The number of SFTSV strains across three different segments. (b) The number of isolates by host category, including humans, companion animals, wild animals, and arthropods (on the left) and the distribution of specific host species among companion and wild animals, and arthropods, respectively (on the right). (c) The number of viral strains by year, from 2009 to 2023 (on the left), and by month (on the right). (d) Geographical distribution of SFTSV strains across China, categorized as Central China regions (Hubei and Henan), Eastern China regions (Anhui, Jiangsu, Zhejiang, Jiangxi, and Taiwan), Northern China regions (Hebei, Beijing, and Shandong), and Northeastern China regions (Liaoning and Jilin), as well as South Korea, Japan, and Thailand. Specific information for each viral strain is given in Supplementary Table S1.

acid similarities from 99.7% to 95.8%. Similarly, S segments exhibited nucleotide similarities from 99.9% to 94.0% and amino acid similarities from 99.8% to 94.8%. These findings suggest a predominantly conservative genetic landscape within SFTSV, indicative of its evolutionary stability across different genomic segments.

Tracing the evolutionary trajectory of SFTSV genotypes through phylogenetic analysis

Utilizing the genotype classification by Fu et al. (2016) and Yun et al. (2020), our analysis of sequence data for each segment of the SFTSV enabled the construction of ML phylogenetic trees (Fu et al. 2016, Yun et al. 2020). This rigorous analysis identified nine distinct genotypes (A, B1, B2, B3, B4, C, D, E, and F) across the L, M, and S genomic segments, as shown in Fig. 2. Notably, the B4 genotype emerged as a consistent presence across all three segments, underscoring its potential significance in the virus's evolution. Predominantly, China primarily reported genotypes A, D, E, and F, with genotype E being unique to this region. In contrast, in South Korea, genotypes within the B series were most prevalent, with B1 and B3 genotypes especially prominent across all segments. Japan observed a higher frequency of genotype B2. The distribution of genotype C was particularly distinctive, appearing predominantly in the L and S segments in Japan and the M segment in China. Additionally, recent isolates from Thailand, although few, have been uniformly identified as belonging to the B3 genotype. This finding underscores the expansion of this genotype beyond its traditional East Asian domain.

Subsequent Bayesian analysis (BEAST) applied to the sequences constructing the ML trees corroborated these results (Fig. 3). While the phylogenetic relationship was similar to those identified through ML methods, the BEAST analysis unveiled distinct evolutionary rates and points of divergence. Specifically, for the L segments, it was estimated that the B genotype diverged from the A, C, D, E, and F genotype clusters around 1610 years, with a 95% HPD ranging from 1571 to 1753 years. Within the B genotype, B4 was the first to branch off around 1660 years (95% HPD = 1630 to 1783), followed by B3 around 1680 years (95% HPD = 1658 to 1793), and finally, B1 and B2 diverged from 1703 years (95% HPD = 1685 to 1816) (Fig. 3a). For the M segments, the B genotypes were posited to have diverged from the other genotypes circa 1010 years (95% HPD = 969 to 1465), with B2 diverging around 1360 years (95% HPD = 1339 to 1654), followed by B4 around 1380 years (95% HPD = 1362 to 1661), and B1 and B3 becoming distinct subgenotypes around 1440 years (95% HPD = 1424 to 1698) (Fig. 3b). Analysis of the S segment sequences similarly suggested that genotype B diverged from the other groups around 1760 years (95% HPD = 1740 to 1873), with B4 splitting first around 1790 years (95% HPD = 1772 to 1866), followed by B2 around 1810 years (95% HPD = 1794 to 1876), and B1 and B3 lastly diverging around 1842 years (95% HPD = 1832 to 1893) (Fig. 3c). These patterns of divergence indicate that the M segment has the slowest evolutionary rate and harbors the oldest most recent common ancestor (tMRCA), whereas the S segment exhibits the fastest evolutionary rate, containing the youngest tMRCA.

Despite the first official report of SFTSV occurring in 2009, the BEAST analysis suggests a much older presence of the virus, marked by genetic variations and segment reassortments across different hosts over time, shaping its evolutionary journey. This reveals a complex and extended history for SFTSV, emphasizing the critical need to delve into its evolutionary dynamics for a deeper understanding of its epidemiology and anticipating potential shifts in the future.

Deciphering the reassortment events of SFTSVs: an in-depth genomic analysis

Next, we focused on unraveling the evolutionary mechanisms of the SFTSV, specifically exploring instances of reassortment within the virus's genome. Utilizing ML phylogenetic trees, we identified a total of 34 potential reassortment genotypes (R1 to R34)

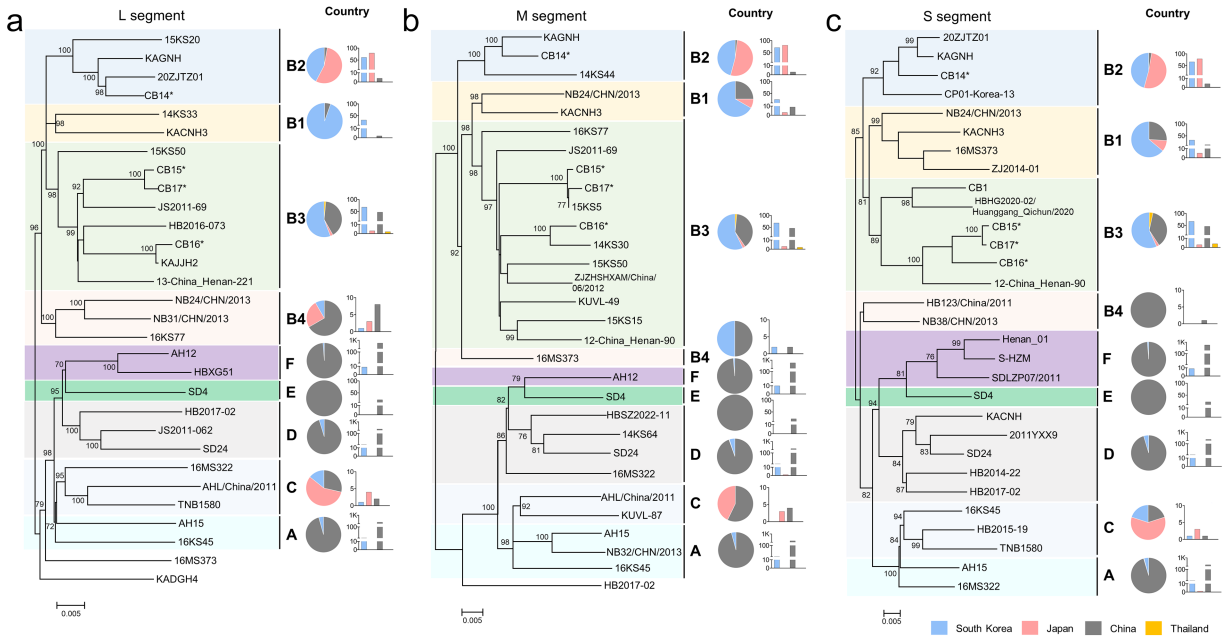

Figure 2. ML phylogenetic trees of the SFTSV genome.

The ML trees were constructed based on the sequence alignment of L (a), M (b), and S (c) segments, respectively. The SFTSV strains detected in this study are indicated by asterisks. The ML trees were tested by bootstrap with 1000 replicates, and only bootstrap values >70% are shown next to the branch nodes. Scale bars indicate nucleotide substitutions per site. Nine genotypes were distinguished against a background of different colors. Pie charts and bar charts next to each phylogenetic tree depict the proportions and numbers of viral strains in different genotypes isolated from South Korea, Japan, China, and Thailand, respectively.

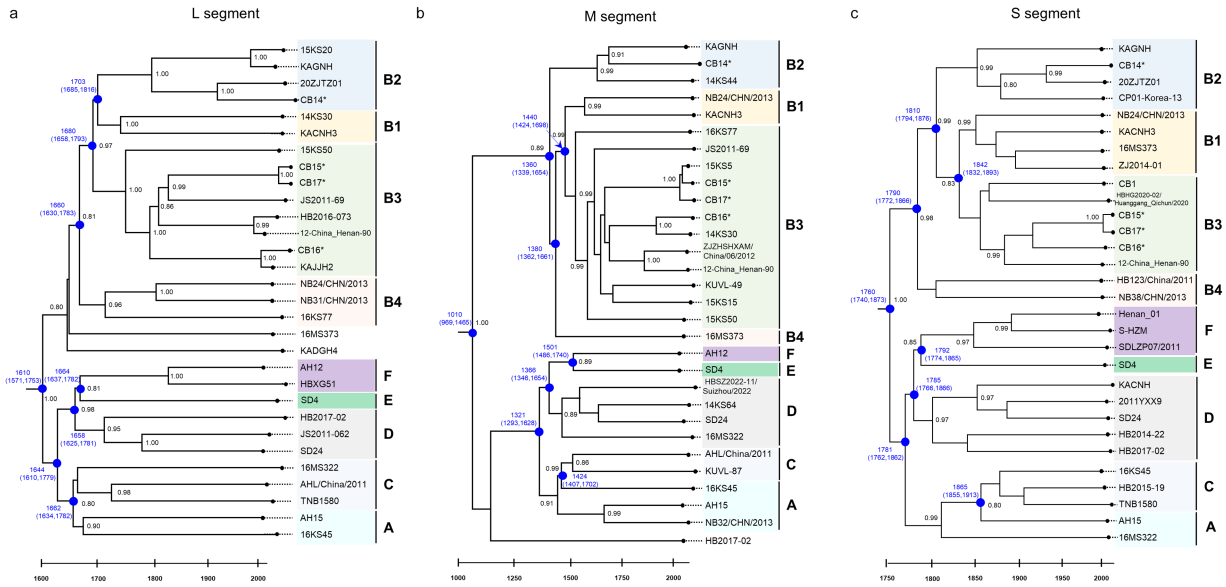

Figure 3. BEAST phylogenetic trees of the SFTSV genome.

The same strains used to infer ML trees were also utilized to infer the MCC tree for the L (a), M (b), and S (c) segments, based on the Bayesian evolutionary inference method. The posterior probability values >0.8 were marked in the evolutionary trees. The estimated median dates for the tMRCA and the corresponding 95% highest posterior density (HPD) of divergence nodes associated with nine SFTSV genotypes (blue dots) were annotated in blue for three different segments.

involving 83 strains (Supplementary Table S2, Fig. 4a). Among these strains, a majority were found in China (60 strains), followed by South Korea (21 strains), and Japan (2 strains). This distribution aligns with the country-specific prevalence of genotype involvement, where reassortments involving genotypes A, D, and F were primarily observed in China. Conversely, in South Korea and Japan, reassortments predominantly involved subtypes of genotype B.

Notably, among the diverse reassortment events, R19 (D/F/D, segment lineage order: L/M/S) was particularly prevalent in China, impacting 10 strains. In South Korea, R11 (B1/B2/B2) was the most commonly observed reassortant, affecting eight strains. Similarly, in China, R8 (B4/B1/B1) and R25 (F/F/D) each affected seven strains. Significantly, six reassortment types—R29 (A/F/D), R30 (B4/A/B1), R31 (C/D/A), R32 (B4/B3/B1), R33 (unclassified (unc)/B2/B1), and R34 (unc/B4/B1)—exhibited unique genotypes across all three genomic segments (L, M, and S), demonstrating extensive genetic diversity. All strains containing the genotype B4 were identified as reassortants, underscoring its consistent presence across all segments. Intriguingly, while most reassortants were isolated from humans, types R8 (B4/B1/B1) and R33 (unc/B2/B1) were identified in a mouse and a cat, respectively, highlighting SFTSV's capacity to infect a broad array of hosts and its adaptability across different species. Furthermore, individual strains of R11 (B1/B2/B2) were discovered in both a dog and a tick (*H. longicornis*), highlighting the complex transmission dynamics of SFTSV and its wide host range.

Tracing the evolution and adaptation of SFTSV across East Asia

Utilizing sequence data categorized by year, country, and host, we meticulously constructed a chart detailing the reassortment dynamics of SFTSV genotypes across China, South Korea, and Japan from 2010 to 2023 (Fig. 4b). In China, the initial reports of genotypes A, D, E, and F date back to 2010, with the B3 genotype emerging in 2011. Of them, genotype F consistently emerged as the most prevalent each year, closely followed by genotypes A and D until 2022 (2023). A noticeable increase in reassortants was observed in 2011, but only R28 (D/F/F), R25 (F/F/D), R19 (D/F/D), and R26 (F/A/F) types were continuously reported until the 2020s. Early detections of segments linked to genotypes C and B2 in China during 2011 and 2012 contrast with the 2013 reports from Japan and South Korea of complete viruses, suggesting a broader dissemination and complex origins for these genotypes. Although the pivotal year of 2013–2014 marked the second emergence of various reassortants in China, only R8 (B4/B1/B1), R22 (F/D/D), and R6 (F/A/A) were noted intermittently up to 2020, suggesting constraints on the sustained presence and spread of new SFTSV reassortants in the natural setting.

In South Korea, after the initial detection of genotype F in 2012, a range of genotypes including B3, B2, A, and D surfaced in 2013, followed by the emergence of B1 in 2014. The year 2016 witnessed genotypes B1, B2, and B3 becoming predominant, producing numerous reassortants. Among these, only R11 (B1/B2/B2) continued to be reported until 2020, indicating a limited period of circulation for novel reassortants. Genotype B3, initially less prevalent than B2, became the most prevalent genotype post-2017 and persisted until 2022 (2023) in South Korea. In Japan, genotype C was reported sparsely until 2017, whereas genotype B2 strains, identified in 2014, were consistently reported up to 2021. Amidst the continuous circulations of major genotypes A, B2, B3, D, E, and F among humans, only a select few reassortants, specifically R25 (F/F/D) and R19 (D/F/D), have been consistently detected through the late 2020s. This observation suggests that only certain reassortant strains have successfully adapted to new hosts and environments, highlighting a selective advantage for specific genomic configurations in the virus's evolution and spread.

In addition to the circulation among humans, SFTSV strains, particularly certain reassortants, have been identified in a diverse range of hosts (Supplementary Table S2). Specifically, R8

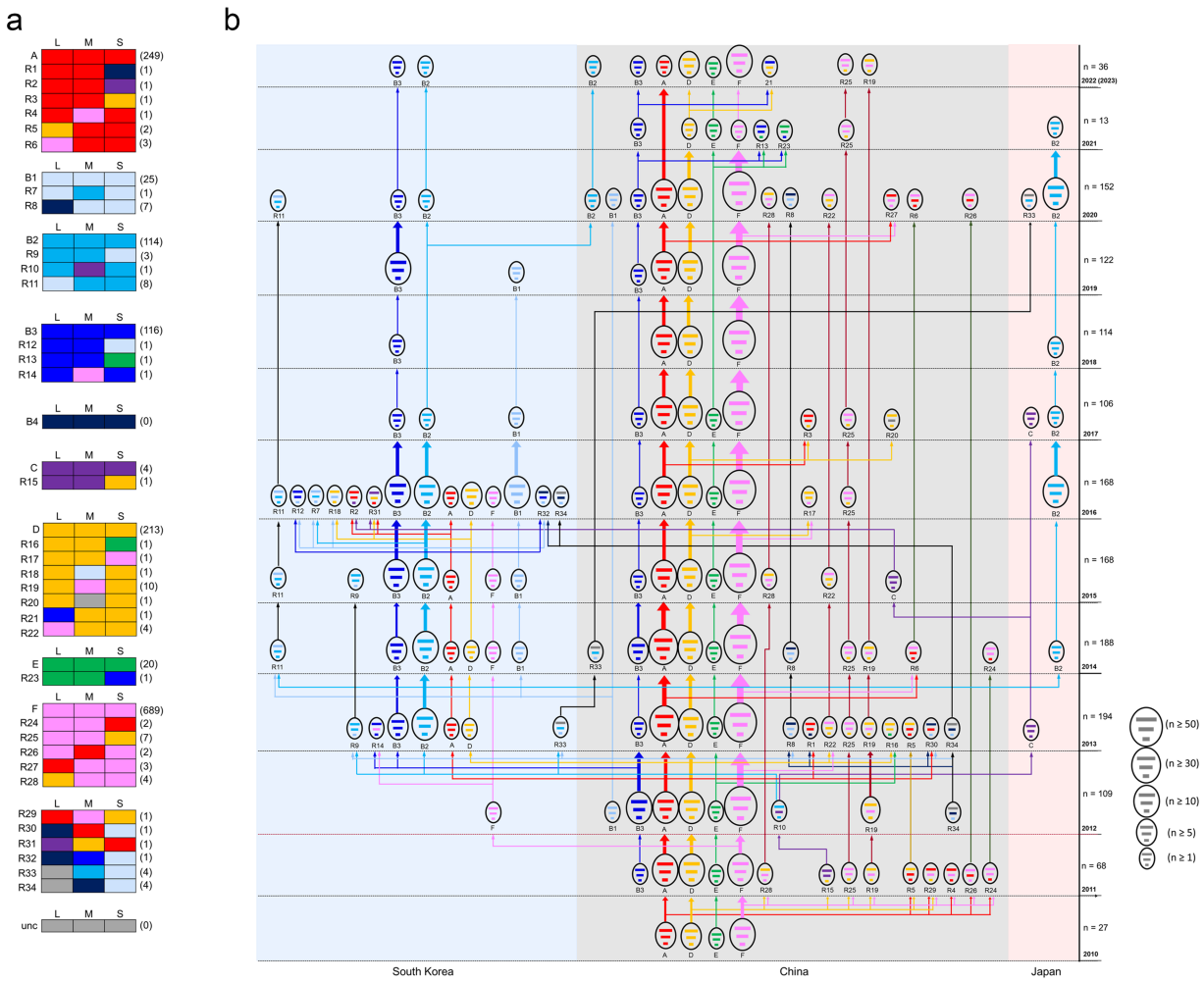

Figure 4. Schematic of the SFTSV evolutionary pathway.

(a) The SFTSVs were assigned to various genotypes depending on the genetic origin of each segment as determined through phylogenetic analysis. They were classified into nine major genotypes (A, B1, B2, B3, B4, C, D, E, and F) and 34 reassortment strains (R1 to R34). In each box, the major genotypes along with its potentially derived reassortment strains were displayed. Each column from the left to right represents the L, M, and S segments, respectively. Colors denote different genotypes: A is red; B1 is pale blue; B2 is cyan; B3 is blue; B4 is navy; C is purple; D is yellow; E is green; F is pink; and unc (uncharacterized) is gray. The number in parentheses to the right of the boxes indicates the count of strains assigned to the corresponding genotypes. Information on reassortment type is listed in Supplementary Table S3. (b) Major genotypes and reassortment strains represented in (a) were charted by different countries and temporal periods from 2010 to 2023. Three countries, South Korea, China, and Japan, are highlighted with backgrounds in blue, gray, and red, respectively. Because the number of strains identified in 2023 was so low, those were combined with the number of strains isolated in 2022 and marked as '2022 (2023)'. The horizontal bars starting from top to bottom in each virion representation indicate L, M, and S segments. The size of each virion and the thickness of each arrow roughly correspond to the number of SFTSV strains of each genotype and reassortment type. The number of viral strains detected each year is noted along the line on the right.

(B4/B1/B1) and R11 (B1/B2/B2) were discovered in a mouse and a dog/tick, respectively, indicating their presence across various animal hosts during prolonged circulation. Building on the analysis, the case of R33 (unc/B2/B1) identified in Japanese cats in 2020 is particularly noteworthy. Given that this reassortant was previously detected in humans in South Korea in 2013 and in China in 2014, but had not been reported in Japan prior to 2020, the appearance of R33 highlights the potential for inter-country and inter-species transmission of the virus. This instance underscores the complex dynamics of SFTSV spread, indicating its ability to traverse geographical and species boundaries. While this analysis is based on GenBank data and may not fully encapsulate the situation in natural and human environments, the reported diversity of virus isolates across countries likely provides a reasonable reflection of SFTSV's current distribution and evolutionary trends.

Deciphering the recombination landscape of SFTSV: an in-depth genomic analysis

To predict recombination events among the virus strains, we employed the RDP5 software package, focusing specifically on each genomic segment. This analysis identified a distribution of recombination events across the L, M, and S segments, with 58 events predicted in the L segment (ReL1 to ReL58), 42 in the M segment (ReM1 to ReM42), and 19 in the S segment (ReS1 to ReS19) (Supplementary Table S3, Fig. 5a). Among the total of 119 recombination events, while most of the events were predicted in Chinese strains (102/119), 17 events were detected across three segments of 16 strains isolated from South Korea. It is noteworthy that several events occurred within the genome segments of the same strain. Briefly, NB32/CHN/2013 strains showed three recombination events in the L segment (ReL3, ReL15, and ReL28) and two in the M segment (ReM10 and ReM27). Moreover, two strains were

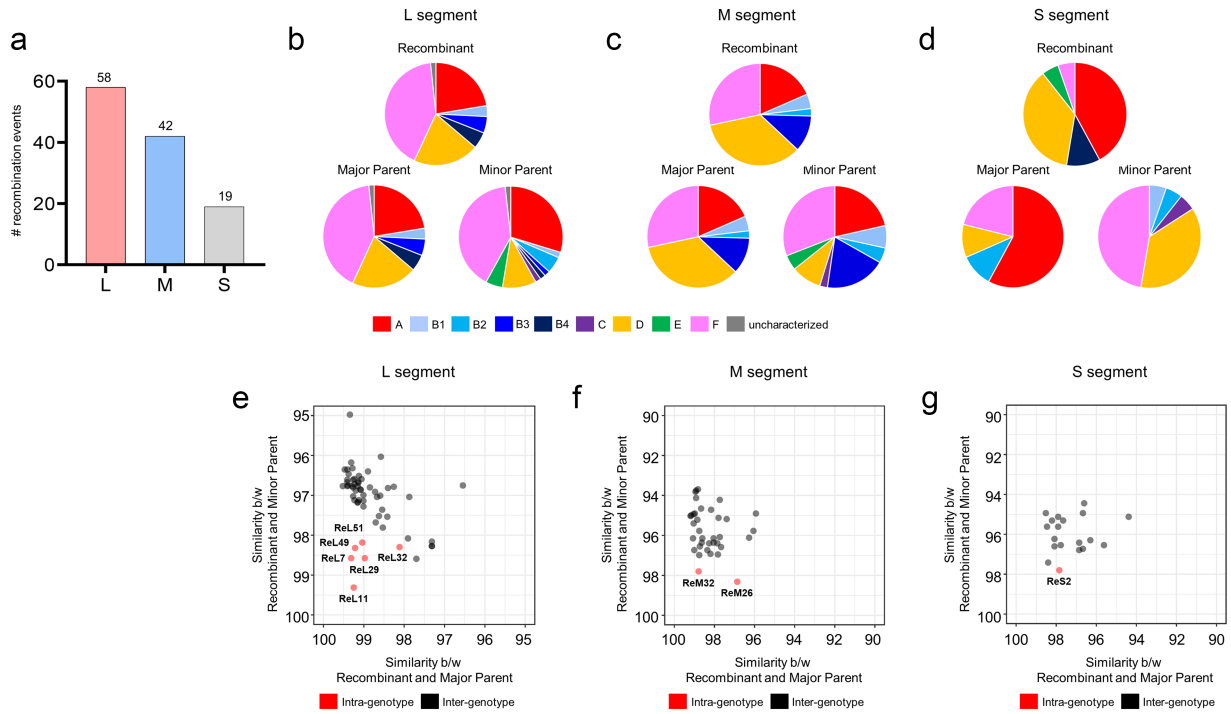

Figure 5. Information on recombination events.

(a) The number of recombination events predicted in different segments. (b-d) The ratio of different genotypes in each recombinant, major parent, and minor parent strain across three segments. (e-g) Scatter plot illustrating the similarity between recombinant and major parent sequences, compared to that between recombinant and minor parent sequences across the L, M, and S segments. While inter-genotype recombination events are marked as black dots in the plots, intra-genotype recombination events are marked as red dots. The names of each intra-genotype recombination event are displayed in the plot.

identified as having three recombination events each: 2011YPQ12 stain, marked by ReL25, ReL51, and ReM16, and 2011YSC60 stain, with ReL26, ReL52, and ReM17. Lastly, nine strains displayed two recombination events (Supplementary Table S3). Interestingly, although most recombinant strains are isolated from humans, recombinant Jilin and SDtickSFTS2018-01 were isolated from ticks, and HB3-sheep03 and SDLZSheep01/2011 strains were identified in sheep.

Analysis of individual recombination events, which categorized major parent and minor parent sequences of the recombinant viruses, revealed that in the L and M segments, the major parent's genotypes were well matched with those of the recombinant strain (Fig. 5b and c). However, the S segment displayed several discrepancies between the major parents and the recombinant strain, indicating a richer genomic diversity within its recombination events compared to the L and M segments (Fig. 5d). This suggests that while the L and M segments primarily undergo short-length gene recombinations, the S segment is characterized by long-length gene recombinations significant enough to alter the genotype from the major parent. Further differentiation of recombination events into inter-genotype and intra-genotype, based on the genotypes of major and minor parents, revealed that most events were inter-genotype, except for six in the L segment (ReL7, ReL11, ReL29, ReL32, ReL49, and ReL51), two in the M segment (ReM26 and ReM32), and one in the S segment (ReS2), which were intra-genotype (Fig. 5e–g).

Identification of hotspots of SFTSV recombination

To identify hotspots where recombination events commonly occurred, we analyzed the start and end points of recombination events across the L, M, and S segments. Briefly, the analysis methodically mapped the start and end points of recombination events within 30-nucleotide (nt) stretches, revealing multiple hotspots across all three segments (Fig. 6a–c). Of the 119 recombination events analyzed, 58 were mapped within the L segment, where start points were identified in 37 stretches and end points in 39, with 9 stretches serving as both start and end points, suggesting that these sequences may facilitate both the initiation and termination of recombination events. In the M segment, recombination initiation and termination were noted in 32 and 24 stretches, respectively, with 4 stretches showing an overlap of start and end points. Similarly, in the S segment, 19 recombination events were observed, with start points in 12 stretches and end points in 7, and 1 stretch featuring both. Notably, termination points were predominantly located at the 3′ end regions of each segment, indicating a possible preference for termination sites. To deepen our understanding, stretches that hosted more than four recombination events were designated as primary hotspots and are marked as ① to ⑨ (Fig. 6a–c). Notably, in the L segment, hotspot ① (3091–3120 nt) was detected prominently in 12 and 1 recombinants as the start and end point, respectively, while the hotspots ② (3691–3720 nt) and ③ (6211–6240 nt) emerged as the recombination termination points, respectively (Fig. 6a). The M segment exhibited hotspots ④ (241–270 nt) and ⑤ (601–630 nt) located at the 5′ end regions, serving as the recombination start ($n = 4$) and end ($n = 4$) points, respectively (Fig. 6b). Notably, hotspot ⑥ (3181–3210 nt) located at the 3′ end region of the M genome disclosed 11 recombinants. Within the S segment, hotspots ⑦ (1051–1080 nt), ⑧ (1291–1320 nt), and ⑨ (1561–1590 nt) were detected, each marked by five events of each region, respectively (Fig. 6c). Remarkably, these hotspots were all located within the NS ORF of the S segment. Notably, the hotspot analysis revealed a high degree of sequence

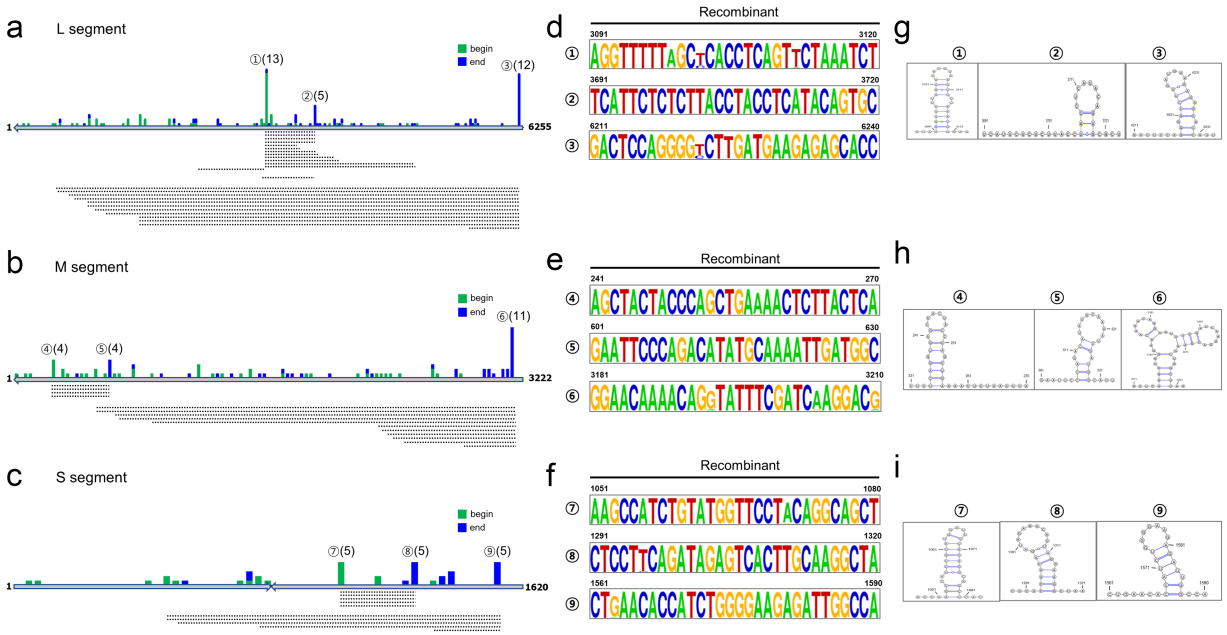

Figure 6. Recombination hotspots.

Based on the stretches divided into 30-nucleotide windows, the start and end positions of each recombination event were charted across three segments: L (a), M (b), and S (c), revealing nine hotspots (① to ⑨). The number in the parentheses next to each hotspot signifies the number of events detected at each spot. The dotted lines below each segment represent the recombined regions associated with the nine hotspots. The degree of conservation was compared among recombinant strains across three segments: L (d), M (e), and S (f). Each comparison involved sequences of 30 nucleotides in length at each spot. Secondary structures were predicted at each spot in L (g), M (h), and S (i) segments, respectively.

conservation among recombinant strains, highlighting the pivotal role of these sequences in the virus's evolutionary adaptability and survival (Fig. 6d–f). The conserved hotspot sequences were further analyzed for RNA secondary structure predictions, revealing a frequent association with various hairpin structures (Fig. 6g–i). Despite varying free energy values across these structures, each stretch involved in more than two recombination events consistently formed hairpin configurations (Supplementary Fig. S1). These distinctive secondary structures are likely crucial for the viral recombination process, facilitating the RdRp enzyme's ability to pause or backtrack, thereby enabling effective template switching.

To determine whether the identified hotspots' recombination events occurred independently across different strains or stemmed from a common ancestral strain, we conducted a comprehensive phylogenetic analysis using BEAST, focusing on recombinant sequences categorized by genotype (Supplementary Fig. S2). For clarity, only genotypes with more than two recombinant strains were included in this analysis. In the L segment, recombination events at identified hotspots (spots ① and ③) involving genotypes A, D, and F were dispersed widely across the phylogenetic tree, suggesting the likelihood of these events occurring independently among the virus populations. Conversely, a notable pattern emerged with recombination events at spot ②, which frequently shared with spot ①, hinting at a potential linkage or common origin between these events (Supplementary Fig. S2A, B and C). Similarly, the recombination events at spots ④ and ⑤ were identified in the same strains within specific genotype D of the phylogenetic tree, indicative of a common ancestral strain. However, the diversity of genotypes involved in recombination events at spot ⑥, including A, D, and F, suggested multiple independent occurrences of recombination across different viruses, underscoring the segment's complex recombination landscape (Supplementary Fig. S2D, E, and F). In the S segment, while some recombination events, especially those related to spot ⑨, seemed to occur independently among recombinants of genotypes A and D, the association of start and end points at spots ⑦ and ⑧ implied the possibility of paired or linked events (Supplementary Fig. S2G and H).

Decoding evolutionary pressures: selection dynamics in SFTSV genes

To understand the evolutionary dynamics of SFTSVs, we investigated the selection pressures exerted on key viral genes. By analyzing the ratio of nonsynonymous to synonymous substitutions (dN/dS) across four essential genes: RdRp, glycoprotein, NP, and NS protein, we discerned a prevailing trend of negative selection. The calculated dN/dS ratios were 0.0657 for RdRp, 0.117 for glycoprotein, 0.0479 for NP, and 0.142 for NS protein, indicating a prevailing trend of negative selection for these genes (Table 1). This trend suggests an evolutionary bias toward conserving these genes, likely attributed to their critical roles in the virus's lifecycle and pathogenicity within various hosts in nature.

Nevertheless, detailed examination identified two residues within the RdRp gene (positions 2 and 1353) that were under strong positive selection, as corroborated by four distinct analytical methods (FEL, FUBAR, SLAC, and MEME). For the glycoprotein and NS genes, each of the four sites (glycoprotein, positions: 170, 298, 501, and 960; NS, positions: 238, 272, 281, and 289) were identified under positive selection residues. The consensus across methods on these sites underscores their likely importance in mediating host cell entry and immune evasion, crucial for viral transmission and survival. However, no positive selection sites were detected in the NP gene, implying that the gene's functions are so vital that any significant alterations could be detrimental to the virus's viability.

Table 1. Molecular selection pressure of the SFTSV genomes.

Gene segment	Coding region	Method	Number of positive selection sites (location of sites)	Number of negative selection sites (%)	dN/dS
L	RdRp	SLAC[a]	4 (2, 719, 1061, 1353)	1461 (70.1)	0.0657
		FEL[b]	5 (2, 211, 691, 1061, 1353)	1700 (81.6)	
		FUBAR[c]	6 (2, 691, 719, 1061, 1116, 1353)	1925 (92.4)	
		MEME[d]	4 (2, 8, 1353, 2063)	n.d.[e]	
M	Glycoprotein	SLAC	6 (13, 37, 170, 298, 501, 960)	688 (64.1)	0.117
		FEL	8 (9, 170, 298, 404, 495, 501, 960, 984)	807 (75.2)	
		FUBAR	6 (9, 170, 298, 501, 960, 1011)	908 (84.6)	
		MEME	7 (298, 404, 495, 501, 524, 642, 960)	n.d.	
S	NP	SLAC	n.d.	166 (67.8)	0.0479
		FEL	n.d.	177 (72.2)	
		FUBAR	n.d.	209 (85.3)	
		MEME	n.d.	n.d.	
	NS	SLAC	6 (145, 238, 249, 272, 281, 289)	207 (70.6)	0.142
		FEL	5 (238, 249, 272, 281, 289)	201 (68.6)	
		FUBAR	7 (17 145, 238, 249, 272, 281, 289)	229 (78.2)	
		MEME	5 (145, 238, 272, 281, 289)	n.d.	

[a]SLAC uses a combination of ML and counting approaches to infer dN and dS substitution rates on a per-site basis. The cutoff *P*-value is ≤.05.
[b]FEL uses an ML approach to infer dN and dS substitution rates on a per-site basis. The cutoff *P*-value is ≤.05.
[c]FUBAR uses a Bayesian approach to infer dN and dS substitution rates on a per-site basis. The cutoff *P*-value is ≥.90.
[d]MEME employs a mixed effects ML approach to test the hypothesis that individual sites have been subject to episodic positive or diversifying selection. The cutoff *P*-value is ≤.05.
[e]n.d., not detected.

Discussion

Our investigation into the evolutionary dynamics of SFTSV across East Asia has yielded critical insights into the virus's genetic diversity, as well as its reassortment and recombination patterns. By analyzing a comprehensive dataset of 2041 sequences through sophisticated methods including ML and BEAST, we have elucidated the complex relationships among nine distinct SFTSV genotypes. This study highlights the pivotal role of genetic diversity in shaping the evolutionary trajectory of SFTSV. The phylogenetic analysis revealed a diverse distribution of SFTSV genotypes across China, South Korea, and Japan, underscoring notable local variations. In South Korea, genotypes B1 and B3 predominate, contrasting with China's broader array of genotypes (A, D, E, and F), which may reflect the influence of regional ecological factors and vector–host interactions in shaping viral genetics. Notably, some strains within the same genotype clustered together across different countries, suggesting the potential for cross-regional transmission of strains. Genotypes B2, B3, and C, initially reported in South Korea and Japan, have since been identified in both China and South Korea, and Japan and China, respectively. Additionally, the detection of genotype B3 in Thailand not only confirms the virus's capability to cross significant geographical barriers but also suggests potential shifts in the epidemiological landscape of SFTSV. Indeed, it has been established that the transnational migration of SFTSV coincides with the East Asia/Australasia flyway used by seabirds traversing South Korea, Japan, and coastal China (Yun et al. 2015; Shi et al. 2017). Moreover, the presence of *H. longicornis* ticks on several migratory bird species (Choi et al. 2014) suggests that migratory birds carrying tick-borne SFTSV could play a crucial role in the virus's cross-national spread. This underscores the need for further research into the interactions between migratory birds, their tick parasites, and SFTSV to better understand and potentially mitigate the transnational transmission of this virus. Future research focusing on the detection and isolation of SFTSV in ticks from migratory birds, along with detailed phylogenetic analysis of these isolates, could enrich our understanding of SFTSV's transmission dynamics.

In this study, we utilized an expansive dataset comprising sequences from 1221 strains from China, 249 strains in South Korea, and 96 strains from Japan. This robust dataset enabled the inference of BEAST phylogenetic trees for three genomic segments, revealing a complex evolutionary pattern of SFTSV that spans centuries, markedly contrasting with its first documented appearance in 2009 (Yu et al. 2011). Notably, the M segment exhibited the slowest evolutionary rate, suggesting that it is a potential locus of genetic stability. Conversely, the S segment, with the youngest most recent common ancestor (tMRCA), underscored areas of rapid evolutionary change. Consistent tree topologies across all segments indicated that genotype B diverged earlier than other genotypes (Fu et al. 2016), and it further differentiated into its subgenotypes, comparable to those observed in genotypes A, C, D, E, and F (Fig. 3), suggesting a relatively long history of virus emergence (Aqian et al. 2021). Furthermore, this study elucidates the reassortment dynamics and evolutionary trajectory of SFTSV across East Asia, highlighting the intricate interplay between genetic diversity, host range, and geographic distribution in shaping the virus's evolution. By identifying 34 distinct reassortment events involving 83 strains across China, South Korea, and Japan, we have uncovered a complex pattern of genetic mixing, characterized by region-specific genotype rearrangements. Genotypes A, D, E, and F are prevalent in China and are involved in the majority of reassortment events, whereas genotypes B1, B2, and B3 and their derivatives dominate in South Korea, suggesting that strategies of localized adaptation are influenced by ecological factors and host interactions. Although a wide array of reassortment types was detected, only a few, such as R8, R19, R22, R25, R26, and R28 in China and R11 in South Korea, demonstrated long-term circulation.

In addition, our analyses extend to examining the relationships between SFTSV fatalities and viral genotypes, revealing that genotype B, which is consistently present in South Korea and Japan, exhibits higher case fatality rates (CFRs) in patients compared to other genotypes (Supplementary Table S4). Notably, genotype B2 shows the highest CFRs among B subtypes, at 47.9%. Genotypes A,

D, and F, predominantly observed in China since 2010, also display relatively high CFRs ranging from 17.3% to 22.7%, in contrast to genotype E's CFR of 12.5% (Supplementary Table S4). This suggests a significant correlation between these specific viral genotypes and elevated mortality rates in these regions. Similar trends are observed with certain reassortment types; e.g., R11, R19, and R25, which are recurrent in South Korea and China, have demonstrated high CFRs in patients (50.0%, 60.0%, and 33.3%, respectively), making them distinct from more transient reassortment types. However, due to the limited number of cases, particularly for reassortants like R18 and R33 which have single reported cases showing a 100% CFR, cautious interpretation is necessary. Thus, ongoing surveillance and additional studies are essential to validate these observations, especially for reassortments reported in fewer than 10 cases. Continued investigation into the genotype-dependent virulence of SFTSV, including detailed monitoring of virus isolations from patients and experimental challenge studies with reassortments through reverse genetic methods, is crucial.

Moreover, our investigation into the recombination landscape of SFTSV across its L, M, and S segments has revealed a complex pattern of genetic exchange crucial to the virus's evolution and adaptation. Despite being a negative-stranded segmented RNA virus, SFTSV exhibits a surprisingly high number of recombination events, suggesting that recombination, alongside reassortment, plays a significant role in its evolutionary dynamics (Chare et al. 2003). Employing the RDP5 software on a large dataset of 2041 strains, we identified a total of 119 recombination events, significantly more than the previous study that used 518 strains and observed fewer events (Aqian et al. 2021). Notably, our study identified recombinant strains such as SDtickSFTS2018-01 and HB3-sheep03 isolated from a tick and a sheep, respectively, adding to the previously reported animal-associated strains like Jilin and SDLZSheep01/2011. Furthermore, our detailed analysis of inter-genotype and intra-genotype recombination events across SFTSV's three genomic segments highlights a predominance of inter-genotype exchanges, suggesting that the genetic variability of SFTSV is extensively driven by cross-genotypic recombination, thereby enhancing the virus's adaptability.

Identifying recombination hotspots across all segments has been instrumental in understanding the nuanced mechanisms of SFTSV evolution. These hotspots reveal a distinct sequence preference at both the commencement and termination of recombination events. Detailed mapping within each segment shows higher sequence conservation among recombinant strains, highlighting the evolutionary importance of these sequences. Further analysis of these hotspots for secondary structures revealed the presence of specific hairpin structures, known to facilitate RdRp pausing, backtracking, and template switching during viral genome recombination (Simon-Loriere et al. 2010, Yang et al. 2021, Bokolia and Gadepalli 2023). The comprehensive mapping of recombination hotspots elucidates the complex interplay of genotypic diversity, underscoring their significance in the virus's evolution. Particularly, the specific locations of these hotspots within the NS ORF of the S segment are of critical importance, marking them as key areas for future research into viral recombination mechanisms and their impact on SFTSV's adaptability and pathogenicity.

Lastly, our selection pressure analysis of the SFTSV genome reveals a predominant trend of negative selection across the RdRp, glycoprotein, NP, and NS genes, underscoring an evolutionary pressure to conserve essential functions critical for the virus's survival and replication. However, the identification of specific sites under positive selection within the RdRp, glycoprotein, and NS genes suggests ongoing evolutionary adaptations, potentially in response to host immune pressures. These findings indicate that these genomic regions are evolving to enhance the virus's ability to evade host defenses and adapt to new hosts or environmental conditions. Notably, the absence of positive selection in the NP gene emphasizes its vital role in the virus's lifecycle, where significant alterations could be detrimental to viral functionality. This analysis illustrates a delicate balance in SFTSV between preserving essential functions through negative selection and engaging in adaptive changes through positive selection. Understanding these evolutionary dynamics is crucial for predicting the virus's future behavior concerning pathogenicity, transmissibility, and resistance to antiviral strategies. Such insights are instrumental in anticipating shifts in the virus's epidemiological profile, thereby informing public health preparedness and response strategies.

While our study provides a comprehensive overview of SFTSV's evolutionary dynamics, it is constrained by the reliance on publicly available GenBank data, which may not capture the full spectrum of viral diversity. Future research should focus on expanding genomic data collection, particularly from underrepresented regions and host species, to refine our understanding of SFTSV's evolutionary patterns. Additionally, functional studies exploring the impact of identified genetic variations on viral fitness and pathogenicity would be invaluable in translating phylogenetic insights into practical strategies for disease control and prevention.

In conclusion, this study significantly enhances our understanding of the genetic and evolutionary dynamics of SFTSV, providing vital insights into its spread and impact on human health. Through integrated phylogenetic and evolutionary analyses, we have not only mapped the complex interplay of genetic diversity and adaptive strategies of SFTSV but also highlighted the importance of these methodologies in forecasting the virus's evolutionary trends. Given the ongoing public health threat posed by SFTSV in East Asia, our findings underscore the urgent need for sustained genetic surveillance and in-depth research into the virus–host–vector interactions, ultimately contributing to improved public health responses and preparedness against SFTSV infections.

Supplementary data

Supplementary data is available at *VEVOLU* online.

Conflict of interest: None declared.

Funding

This work was supported by the Institute for Basic Science (no. IBS-R801-D1), by the National Research Foundation of Korea (no. NRF-2020R1A2C3008339), and by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry through the Pet Industry Technology Development Program (no. RS-2022-IP322088). This work was also supported by the National Institutes of Health (nos. AI171201 and AI152190).

Data availability

Genome sequences and annotations of the four SFTSVs identified in this study have been deposited to the GenBank database (https://www.ncbi.nlm.nih.gov/nucleotide/) under the accession numbers PP656910–PP656921.

References

Aqian L, Liu L, Wu W *et al*. Molecular evolution and genetic diversity analysis of SFTS virus based on next-generation sequencing. *Biosaf Health* 2021;**3**:105–15.

Bokolia NP, Gadepalli R. Identification of consensus hairpin loop structure among the negative sense subgenomic RNAs of SARS-CoV-2. *Bull Natl Res Cent* 2023;**47**:28.

Casel MA, Park SJ, Choi YK. Severe fever with thrombocytopenia syndrome virus: emerging novel phlebovirus and their control strategy. *Exp Mol Med* 2021;**53**:713–22.

Chare ER, Gould EA, Holmes EC. Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. *J Gen Virol* 2003;**84**:2691–703.

Choi C-Y, Kang C-W, Kim E-M *et al*. Ticks collected from migratory birds, including a new record of *Haemaphysalis formosensis*, on Jeju Island, Korea. *Exp Appl Acarol* 2014;**62**:557–66.

Darty K, Denise A, Ponty Y. VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics* 2009;**25**:1974–75.

Fu L, Niu B, Zhu Z *et al*. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 2012;**28**:3150–52.

Fu Y, Li S, Zhang Z *et al*. Phylogeographic analysis of severe fever with thrombocytopenia syndrome virus from Zhoushan Islands, China: implication for transmission across the ocean. *Sci Rep* 2016;**6**:19563.

Gai Z, Liang M, Zhang Y *et al*. Person-to-person transmission of severe fever with thrombocytopenia syndrome bunyavirus through blood contact. *Clin Infect Dis* 2012;**54**:249–52.

Kim J, Hong H-J, Hwang J-H *et al*. Risk factors associated with death due to severe fever with thrombocytopenia syndrome in hospitalized Korean patients (2018–2022). *Osong Public Health Res Perspect* 2023;**14**:151–63.

Kosakovsky Pond SL, Poon AFY, Velazquez R *et al*. HyPhy 2.5—a customizable platform for evolutionary hypothesis testing using phylogenies. *Mol Biol Evol* 2020;**37**:295–99.

Liu B, Zhu J, He T *et al*. Genetic variants of *Dabie bandavirus*: classification and biological/clinical implications. *Virol J* 2023;**20**:68.

Madeira F, Pearce M, Tivey ARN *et al*. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res* 2022;**50**:W276–79.

Martin DP, Varsani A, Roumagnac P *et al*. RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. *Virus Evol* 2021;**7**:veaa087.

McDonald SM, Nelson MI, Turner PE *et al*. Reassortment in segmented RNA viruses: mechanisms and outcomes. *Nat Rev Microbiol* 2016;**14**:448–60.

Nguyen L-T, Schmidt HA, von Haeseler A *et al*. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;**32**:268–74.

Perez-Losada M, Arenas M, Galán JC *et al*. Recombination in viruses: mechanisms, methods of study, and evolutionary consequences. *Infect Genet Evol* 2015;**30**:296–307.

Proctor JR, Meyer IM. CoFold: an RNA secondary structure prediction method that takes co-transcriptional folding into account. *Nucleic Acids Res* 2013;**41**:e102.

Rozewicki J, Li S, Amada KM *et al*. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res* 2019;**47**:W5–10.

Sasaya T, Palacios G, Briese T *et al*. ICTV virus taxonomy profile: phenuiviridae 2023. *J Gen Virol* 2023;**104**:001893.

Shi J, Hu S, Liu X *et al*. Migration, recombination, and reassortment are involved in the evolution of severe fever with thrombocytopenia syndrome bunyavirus. *Infect Genet Evol* 2017;**47**:109–17.

Simon-Loriere E, Martin DP, Weeks KM *et al*. RNA structures facilitate recombination-mediated gene swapping in HIV-1. *J Virol* 2010;**84**:12675–82.

Suchard MA, Lemey P, Baele G *et al*. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol* 2018;**4**:vey016.

Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol* 2021;**38**:3022–27.

Tran XC, Yun Y, Van An L *et al*. Endemic severe fever with thrombocytopenia syndrome, Vietnam. *Emerg Infect Dis* 2019;**25**:1029–31.

Wang Y, Pang B, Wang Z *et al*. Genomic diversity and evolution analysis of severe fever with thrombocytopenia syndrome in East Asia from 2010 to 2022. *Front Microbiol* 2023;**14**:1233693.

Win AM, Nguyen YTH, Kim Y *et al*. Genotypic heterogeneity of *Orientia tsutsugamushi* in scrub typhus patients and thrombocytopenia syndrome co-infection, Myanmar. *Emerg Infect Dis* 2020;**26**:1878–81.

Wu Z, Han Y, Liu B *et al*. Decoding the RNA viromes in rodent lungs provides new insight into the origin and evolutionary patterns of rodent-borne pathogens in Mainland Southeast Asia. *Microbiome* 2021;**9**:18.

Yang Y, Yan W, Hall AB *et al*. Characterizing transcriptional regulatory sequences in coronaviruses and their role in recombination. *Mol Biol Evol* 2021;**38**:1241–48.

Yu X-J, Liang M-F, Zhang S-Y *et al*. Fever with thrombocytopenia associated with a novel bunyavirus in China. *N Engl J Med* 2011;**364**:1523–32.

Yun S-M, Park S-J, Kim Y-I *et al*. Genetic and pathogenic diversity of severe fever with thrombocytopenia syndrome virus (SFTSV) in South Korea. *JCI Insight* 2020;**5**:e129531.

Yun Y, Ryu SY, Heo ST *et al*. Phylogenetic analysis of severe fever with thrombocytopenia syndrome virus in South Korea and migratory bird routes between China, South Korea, and Japan. *Am J Trop Med Hyg* 2015;**93**:468–74.

Zhuang L, Sun Y, Cui X-M *et al*. Transmission of severe fever with thrombocytopenia syndrome virus by *Haemaphysalis longicornis* ticks, China. *Emerg Infect Dis* 2018;**24**:868–71.

Virus Evolution, 2024, 10(1), veae054 , DOI: https://doi.org/10.1093/ve/veae054, Advance Access Publication 20 July 2024Research Article