Gene Reports 34 (2024) 101875

Contents lists available at ScienceDirect

Gene Reports

journal homepage: www.elsevier.com/locate/genrep

Classification of sapoviruses based on comparison of phylogenetic trees for structural and non-structural proteins

Yoshiyuki Suzuki [a,*], Hirokazu Kimura [b], Kazuhiko Katayama [c]

[a] Graduate School of Science, Nagoya City University, 1 Yamanohata, Nagoya-shi, Aichi-ken 467-8501, Japan
[b] Graduate School of Health Science, Gunma Paz University, 1-7-1 Tonyamachi, Takasaki-shi, Gunma-ken 370-0006, Japan
[c] Omura Satoshi Memorial Institute, Graduate School of Infection Control Sciences, Kitasato University, 5-9-1 Shirokane, Minato-ku, Tokyo-to 108-8641, Japan

ARTICLE INFO

Edited by Jormay Lim

Keywords:
Genogroup
Phylogenetic tree
Recombination
Sapovirus
Shared cluster

ABSTRACT

Sapoviruses (SaVs) compose the *Sapovirus* genus in the *Caliciviridae* family. In the genome of SaVs, the junction of coding regions for the non-structural proteins (NSs) and the major structural protein (viral protein 1: VP1) is known as the recombination hotspot. SaVs have been divided into genogroups GI–GXIX based on the phylogenetic relationship of amino acid sequences for VP1. Here, phylogenetic analyses were conducted for NSs and VP1 of 326 SaV strains retrieved from the International Nucleotide Sequence Database. Shared clusters were identified between the phylogenetic trees, and genogroups of SaVs were classified into 10 shared cluster groups, which may represent the monophyletic recombining groups of SaVs. All of human GI, GII, GIV, and GV strains were classified into the same shared cluster group. Continuous surveillance may be warranted for monitoring emergence of intra-genogroup and inter-genogroup recombinants in human SaVs.

1. Introduction

Sapoviruses (SaVs) infect mammals, and cause acute gastroenteritis (AGE) in humans and pigs of all ages (Madeley and Cosgrove, 1976). In humans, SaVs are responsible for ~5 % of AGE among children <5 years of age and becoming more prevalent after the introduction of rotavirus vaccines (Diez Valcarce et al., 2021; Okitsu et al., 2021; Hoque et al., 2022). SaVs compose the *Sapovirus* genus in the *Caliciviridae* family (Oka et al., 2015). The genome of SaV is a single-stranded, positive-sense RNA of 7.1–7.7 kb, mainly containing ORF1 and ORF2. ORF1 encodes a polyprotein, which is cleaved into six non-structural proteins (NSs), NS1–NS5 and NS6/NS7, and the major structural protein, viral protein 1 (VP1). ORF2 encodes the minor structural protein, VP2. ORF3 is occasionally located within the coding region for VP1, encoding a protein with unknown function (Liu et al., 1995).

SaVs mainly infect villous enterocytes of proximal small intestine through binding of VP1 to sialic acids and occludin, which serve as the attachment factor and the co-receptor, respectively, with the aid of claudin-1, which serves as the entry factor (Alfajaro et al., 2019). VP1 also determines the antigenicity of SaVs. Based on the phylogenetic relationship of amino acid sequences for VP1, SaVs have been divided into genogroups GI–GXIX (Oka et al., 2016; Yinda et al., 2017). GI, GII, GIV, and GV have been identified in humans. Various genogroups have also been identified in other mammals, including pigs (GIII and GV–GXI), sea lions (GV), mink (GXII), dogs (GXIII), bats (GXIV and GXVI–GXIX), and rats (GXV). Genogroups have been further divided into at least 52 genotypes (Li et al., 2018; Diez Valcarce et al., 2019).

During the replication cycle of SaVs within the cytoplasm, subgenomic RNAs containing coding regions for VP1 and VP2 are generated in addition to genomic RNAs (Chang et al., 2004). The junction of coding regions for NSs and VP1 appears to contain the promoter for subgenomic RNAs with similar nucleotide sequence and secondary structure as the promoter for genomic RNAs in the negative-sense RNA (Simmonds et al., 2008). Then, recombination may be induced through template switching at the junction of coding regions for NSs and VP1, which is known as the recombination hotspot (Kuroda et al., 2017). Occurrences of intra-genogroup and inter-genogroup recombinations upon co-infection of the same cells during evolution of SaVs have been inferred from the phylogenetic analyses, because recombinants may be located at different places in the phylogenetic trees constructed for the 5′

Abbreviations: AGE, acute gastroenteritis; AICc, corrected Akaike information criterion; BIC, Bayesian information criterion; F, equilibrium frequencies of amino acids as estimated from data; Γ, gamma distribution; GNA, genogroup not assigned; I, invariable sites; INSD, International Nucleotide Sequence Database; LG, Le-Gascuel; ML, maximum likelihood; NJ, neighbor-joining; NoV, norovirus; NS, non-structural protein; SaV, sapovirus; VP, viral protein.

* Corresponding author.
E-mail address: yossuzuk@nsc.nagoya-cu.ac.jp (Y. Suzuki).

https://doi.org/10.1016/j.genrep.2023.101875
Received 9 October 2023; Received in revised form 27 November 2023; Accepted 19 December 2023
Available online 23 December 2023

and 3′ regions of recombination hotspot. Intra-genogroup recombinations have been identified within GI, GII, and GIII (Katayama et al., 2004; Wang et al., 2005; Wang et al., 2020), whereas inter-genogroup recombinations have been identified between GII and GIV, GVII and GXI, and GIX and GX (Hansman et al., 2005; Chanit et al., 2009; Oka et al., 2015; Kuroda et al., 2017).

Genogroups reflect the phylogenetic relationship for the 3′ region of recombination hotspot. However, as the relationship may vary between the 5′ and 3′ regions, both regions should be utilized for classifying SaVs. In noroviruses (NoVs), which possess similar genomic organization and recombination hotspot as SaVs, the 5′ and 3′ regions have been independently classified into P-groups and genogroups, respectively (Katayama et al., 2002; Kroneman et al., 2013; Chhabra et al., 2019). Each NoV strain is characterized by a combination of P-group and genogroup, and thus the number of taxonomic groups increases when recombinants with novel combinations of P-group and genogroup are discovered. In contrast, the biological species concept asserts that individuals exchanging genetic materials should be classified into the same group (Bobay and Ochman, 2018). In this case, the number of taxonomic groups decreases when recombinants between different groups are newly discovered, because these groups would be merged. However, since recombinants may not necessarily be produced between monophyletic groups, there may be a possibility of merging paraphyletic groups.

To reconcile classifications based on the phylogenetic relationship and the exchangeability of genetic materials, SaVs may be classified into the monophyletic recombining groups; i.e., monophyletic groups of strains that may recombine with strains from the same group but not with those from different groups. Such groups may form distinct clusters in both the phylogenetic trees for the 5′ and 3′ regions of recombination hotspot, called the shared clusters, although the phylogenetic relationships among and within the groups may vary between the regions due to occurrences of recombinations in the ancestral population and each group, respectively (Suzuki et al., 2023). The purpose of the present study was to classify SaVs based on the shared clusters.

2. Materials and methods

2.1. Sequence data

Genomic sequences for SaV strains ≥5 kb were retrieved from the International Nucleotide Sequence Database (INSD) on June 24, 2023. Nucleotide sequences for ORF1 were translated into amino acid sequences. Sequences were eliminated if they were identical to other sequences or contained ≥5 % ambiguous amino acids. The remaining 326 sequences were separated into the regions of NSs and VP1, which represented the 5′ and 3′ regions of recombination hotspot, respectively (Supplementary Table S1). The numbers of amino acid sites included in 326 sequences ranged from 1051 to 1763 for the 5′ region and from 529 to 571 for the 3′ region.

2.2. Data analysis

Multiple alignments of 326 amino acid sequences for the 5′ and 3′ regions of recombination hotspot were made using the computer program MAFFT (version 7.247) (Katoh et al., 2002). Amino acid positions containing gaps were eliminated from the alignments for the 5′ and 3′ regions, which finally included 896 sites and 406 sites, respectively. The best fit models of amino acid substitutions for the 5′ and 3′ regions were selected by the corrected Akaike information criterion (AICc) and the Bayesian information criterion (BIC) using MEGA (version 10.1.7) (Kumar et al., 2018). The same model was selected for the 5′ and 3′ regions by AICc and BIC; the Le-Gascuel (LG) model (Le and Gascuel, 2008) with gamma-distributed rate-heterogeneity among sites (Γ) including invariable sites (I) assuming equilibrium frequencies of amino acids as estimated from the data (F) (LG + Γ + I + F).

The maximum likelihood (ML) method with the LG + Γ + I + F model was adopted to construct phylogenetic trees of 326 amino acid sequences for the 5′ and 3′ regions of recombination hotspot using MEGA (version 10.1.7) (Kumar et al., 2018). The neighbor-joining (NJ) method with the p distance, which has been reported to produce more reliable topologies than the ML method (Nei and Kumar, 2000; Yoshida and Nei, 2016), was also adopted using MEGA (version 10.1.7) (Kumar et al., 2018). Reliabilities of interior branches in the phylogenetic trees were examined by the bootstrap method with 1000 re-samplings. Phylogenetic trees were drawn using FigTree (version 1.4.4) (Rambaut, 2018).

Shared clusters were identified as the clusters made by the same sets of SaV strains in the phylogenetic trees for the 5′ and 3′ regions of recombination hotspot. SaV strains were classified into as many shared clusters as possible, which were called the shared cluster groups. The shared cluster groups may represent the monophyletic recombining groups of SaVs (Suzuki et al., 2023).

3. Results

3.1. Phylogenetic trees

SaVs have been divided into genogroups GI-GXIX based on the phylogenetic relationship of amino acid sequences for VP1, which represents the 3′ region of recombination hotspot (Oka et al., 2016; Yinda et al., 2017). Thus, the ML tree of 326 SaV strains was first constructed for the 3′ region (Supplementary Fig. S1). Although no GIX and only one GXIII strain was contained in the data (Nagai et al., 2020), SaV strains belonging to each of other genogroups formed an individual cluster, consistent with the genogroup classification (Oka et al., 2016; Yinda et al., 2017). Two rat and two bat strains that have not been assigned to GI–GXIX formed individual clusters distinct from GI–GXIX, and were tentatively designated as genogroups not assigned 1 (GNA1) and GNA2, respectively (Supplementary Table S1). Similar clustering patterns were also observed in the NJ tree constructed for the 3′ region (Supplementary Fig. S2).

When the ML tree of 326 SaV strains was constructed for the 5′ region of recombination hotspot, GIV strains were included in the cluster of GII strains (Supplementary Fig. S3). GX and GXI strains were also included in the cluster of GVII strains. Two GXV strains did not form a cluster but diverged successively from the branch leading to the cluster composed of GVI, GVII, GX, and GXI strains. GXVIII and GXIX strains formed a mixed cluster. In addition, sea lion GV strain was separated from the cluster of human and pig GV strains. These observations were largely consistent with the previous findings that inter-genogroup recombinations occurred between GII and GIV (Hansman et al., 2005; Chanit et al., 2009) and GVII and GXI (Oka et al., 2016; Kuroda et al., 2017) and that sea lion GV strain was an inter-genogroup recombinant (Oka et al., 2015). Similar clustering patterns were also observed in the NJ tree constructed for the 5′ region (Supplementary Fig. S4).

3.2. Shared clusters

There were 93 shared clusters identified between the ML trees constructed for the 5′ and 3′ regions of recombination hotspot (Table 1; Supplementary Figs. S1 and S3). Genogroups of SaVs were classified into 10 shared cluster groups (Table 2). All of human GI, GII, GIV, and GV strains as well as pig and sea lion GV strains constituted a shared cluster group together with pig GVIII and dog GXIII strains. Pig GVI, GVII, GX, and GXI strains constituted a shared cluster group together with rat GXV strains. Reportedly, GVI, GVII, GX, and GXI strains demonstrated distinctive features in secondary structures at the 5′ end of the genome and the starting region of the recombination hotspot, which may facilitate recombinations among these genogroups (Sunaga et al., 2019). Although pig GIX strains were not contained in the data, these strains have been clustered with GVI, GX, and GXI strains in the phylogenetic analyses for VP1, NS7, and VP2, respectively (Nagai et al., 2020). In

Table 1
Relationship among genogroups of SaVs based on shared clusters.

ML[a]	NJ[b]	Relationship[c]
93	92	{{{[{GI}, GII, {GIV}, GV, {GVIII}, GXIII], [GXII]}, [GIII]}, {[{{GVI}, GVII, {GX}, {GXI}}, GXV], [GNA1]}, {{[GXIV], [GXVI]}, [GXVII]}, [{GXVIII}, GXIX], [GNA2]}[d]

[a] Number of shared clusters identified in the comparison of ML trees for the 5′ and 3′ regions of recombination hotspot.
[b] Number of shared clusters identified in the comparison of NJ trees for the 5′ and 3′ regions of recombination hotspot.
[c] The same relationship was observed in the comparisons of ML trees and NJ trees.
[d] Genogroups included in shared cluster groups are listed in bold square brackets, whereas those included in other shared clusters are listed in plain curly brackets.

Table 2
Genogroups of SaVs in shared cluster groups.

Genogroup[a]	Host[c]	3′ region		5′ region	
		Bootstrap probabiity[d]	Lowest identity[e]	Bootstrap probabiity[f]	Lowest identity[g]
GI, GII, GIV, GV, GVIII, GXIII	Humans, pigs, dogs, sea lions	79.8, 74.5	0.453	100.0, 100.0	0.540
GIII	Pigs	100.0, 100.0	0.783	100.0, 100.0	0.814
GVI, GVII, (GIX)[b], GX, GXI, GXV	Pigs, rats	93.0, 91.3	0.463	99.7, 99.6	0.429
GXII	Mink	100.0, 100.0	0.904	100.0, 100.0	0.958
GXIV	Bats	100.0, 100.0	0.993	100.0, 100.0	0.997
GXVI	Bats	100.0, 100.0	0.719	100.0, 100.0	0.661
GXVII	Bats	100.0, 100.0	0.761	100.0, 100.0	0.722
GXVIII, GXIX	Bats	100.0, 100.0	0.534	100.0, 100.0	0.606
GNA1	Rats	100.0, 100.0	0.988	100.0, 100.0	0.994
GNA2	Bats	99.9, 100.0	0.522	100.0, 100.0	0.619

[a] Genogroups included in each shared cluster group are listed in each row.
[b] Although GIX strains were not contained in the data, these strains were considered to be included in the shared cluster group of GVI, GVII, GX, GXI, and GXV strains.
[c] Hosts for genogroups of SaVs included in each shared cluster group are listed in each row.
[d] Bootstrap probabilities supporting the shared cluster group in the ML (left) and NJ (right) trees for the 3′ region of recombination hotspot.
[e] The lowest pairwise amino acid sequence identity within the shared cluster group for the 3′ region of recombination hotspot.
[f] Bootstrap probabilities supporting the shared cluster group in the ML (left) and NJ (right) trees for the 5′ region of recombination hotspot.
[g] The lowest pairwise amino acid sequence identity within the shared cluster group for the 5′ region of recombination hotspot.

addition, GIX strains appeared to have produced recombinants with GX strains (Kuroda et al., 2017). These observations suggested that GIX strains may be included in the shared cluster group of GVI, GVII, GX, GXI, and GXV strains. Bat GXVIII and GXIX strains also constituted a shared cluster group. Each of other genogroups, namely pig GIII strains, mink GXII strains, bat GXIV, GXVI, GXVII, and GNA2 strains, and rat GNA1 strains, constituted an individual shared cluster group. All shared cluster groups were supported with relatively high bootstrap probabilities, especially in the phylogenetic tree constructed for the 5′ region, which included a greater number of amino acid sites than the 3′ region (Table 2; Supplementary Figs. S1 and S3). Similar number (92) of shared clusters as well as the same set of 10 shared cluster groups supported with similarly high bootstrap probabilities were also observed in the comparison of NJ trees constructed for the 5′ and 3′ regions (Tables 1 and 2; Supplementary Figs. S2 and S4). The lowest pairwise amino acid sequence identity within each shared cluster group ranged from 0.453 to 0.993 for the 3′ region and from 0.429 to 0.997 for the 5′ region (Table 2). On the other hand, the pairwise identity between shared cluster groups ranged from 0.305 to 0.557 for the 3′ region and from 0.280 to 0.526 for the 5′ region.

4. Discussion

4.1. Classification of SaVs based on shared clusters

In the present study, SaVs were classified taking into account the phylogenetic relationship and the exchangeability of genetic materials. The clustering patterns in the phylogenetic trees of 326 SaV strains constructed for the 3′ region of recombination hotspot were consistent with the genogroup classification (Oka et al., 2016; Yinda et al., 2017), and those constructed for the 5′ region were largely consistent with the previous findings for inter-genogroup recombinations (Hansman et al., 2005; Chanit et al., 2009; Oka et al., 2015; Oka et al., 2016; Kuroda et al., 2017). Shared clusters were identified between the phylogenetic trees for the 5′ and 3′ regions, and genogroups of SaVs were classified into 10 shared cluster groups. Since the shared cluster groups were defined so as to satisfy the biological species concept (Bobay and Ochman, 2018), they may be candidates for species in the *Sapovirus* genus. Notably, the lowest pairwise amino acid sequence identity within shared cluster groups was found to be smaller than the highest identity between shared cluster groups, suggesting that shared cluster groups may not be identified using cut-off values in the amino acid sequence identity. In practice, however, shared cluster groups were identified as the groups of strains without evidence for occurrences of recombinations with strains from different groups. Thus, some of the shared cluster groups identified in the present study may be divided or merged as more SaV strains are sampled for determining genomic sequences and included in the analysis.

4.2. Implications from shared cluster groups

The shared cluster groups may represent the monophyletic recombining groups of SaVs (Suzuki et al., 2023). For the different shared cluster groups infecting different hosts, e.g., pig GIII and mink GXII strains, suppression of co-infection due to differentiation of host tropisms may prevent recombinations. On the other hand, two, three, and five shared cluster groups appeared to co-circulate in rats, pigs, and bats, respectively, suggesting that the 5′ and 3′ regions of recombination hotspot may be incompatible between different groups within these hosts. In particular, bats appeared to harbor the largest number of shared cluster groups. Notably, bats are known to exhibit a unique ability to host diverse viruses without suffering from diseases, which is considered to be conferred by the balance between enhanced defense responses and immune tolerance (Irving et al., 2021).

Although intra-genogroup recombinations within GI and GII (Katayama et al., 2004) and inter-genogroup recombinations between GII and GIV (Hansman et al., 2005; Chanit et al., 2009) have been identified in human SaVs, it has also been reported that recombination frequency may be low in human SaVs compared to human NoVs (Tohma et al., 2020; Doan et al., 2023). This may be caused by lower prevalence and thus less opportunity for co-infection in human SaVs than in human

NoVs (Hassan et al., 2019). In addition, nucleotide sequence at the recombination hotspot in the SaV genome appeared to be less conserved than that in the NoV genome, which may limit template switching in SaVs compared to NoVs (Tohma et al., 2020). Differences in genomic organizations, e.g., NSs and VP1 are encoded by ORF1 and ORF2, respectively, in the NoV genome, whereas these proteins are encoded by ORF1 in the SaV genome, may also contribute to less recombinations in SaVs than in NoVs (Liu et al., 1995). In the present study, however, all of human GI, GII, GIV, and GV strains were classified into the same shared cluster group, implying that recombinants may be produced between these genogroups (Hansman et al., 2005; Chanit et al., 2009). Furthermore, prevalence of SaVs appears to be increasing in humans (Diez Valcarce et al., 2021; Okitsu et al., 2021; Hoque et al., 2022). It has also been reported that recombinations may be an adaptive mechanism in NoVs (Suzuki, 2021). Continuous surveillance may thus be warranted for monitoring emergence of intra-genogroup and inter-genogroup recombinants in human SaVs.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.genrep.2023.101875.

CRediT authorship contribution statement

Yoshiyuki Suzuki: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft. **Hirokazu Kimura:** Writing – review & editing, Conceptualization, Validation. **Kazuhiko Katayama:** Writing – review & editing, Conceptualization, Validation.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP19K12221 and AMED Grant Number JP23fk0108667 to Y.S. The authors thank anonymous reviewers for valuable comments.

References

Alfajaro, M.M., Cho, E.-H., Kim, D.-S., Kim, J.-Y., Park, J.-G., Soliman, M., Baek, Y.-B., Park, C.-H., Kang, M.-I., Park, S.-I., Cho, K.-O., 2019. Early porcine sapovirus infection disrupts tight junctions and uses occludin as a coreceptor. J. Virol. 93 e01773-18.

Bobay, L.-M., Ochman, H., 2018. Biological species in the viral world. Proc. Natl. Acad. Sci. U. S. A. 115, 6040–6045.

Chang, K.O., Sosnovtsev, S.V., Belliot, G., Kim, Y., Saif, L.J., Green, K.Y., 2004. Bile acids are essential for porcine enteric calicivirus replication in association with down-regulation of signal transducer and activator of transcription 1. Proc. Natl. Acad. Sci. U. S. A. 101, 8733–8738.

Chanit, W., Thongprachum, A., Khamrin, P., Okitsu, S., Mizuguchi, M., Ushijima, H., 2009. Intergenogroup recombinant sapovirus in Japan, 2007–2008. Emerg. Infect. Dis. 15, 1084–1087.

Chhabra, P., de Graaf, M., Parra, G.I., Chan, M.C.-W., Green, K., Martella, V., Wang, Q., White, P.A., Katayama, K., Vennema, H., Koopmans, M.P.G., Vinje, J., 2019. Updated classification of norovirus genogroups and genotypes. J. Gen. Virol. 100, 1393–1406.

Diez Valcarce, M., Montmayeur, A., Tatusov, R., Vinje, J., 2019. Near-complete human sapovirus genome sequences from Kenya. Microbiol. Resour. Announc. 8 e01602-18.

Diez Valcarce, M., Kambhampati, A.K., Calderwood, L.E., Hall, A.J., Mirza, S.A., Vinje, J., 2021. Global distribution of sporadic sapovirus infections: a systematic review and meta-analysis. PloS One 16, e0255436.

Doan, Y.H., Yamashita, Y., Shinomiya, H., Motoya, T., Sakon, N., Suzuki, R., Shimizu, H., Shigemoto, N., Harada, S., Yahiro, S., Tomioka, K., Sakagami, A., Ueki, Y., Komagome, R., Saka, K., Okamoto-Nakagawa, R., Shirabe, K., Mizukoshi, F., Arita, Y., Haga, K., Katayama, K., Kimura, H., Muramatsu, M., Oka, T., 2023. Distribution of human sapovirus strain genotypes over the last four decades in Japan: a global perspective. Jpn. J. Infect. Dis. 76, 255–258.

Hansman, G.S., Takeda, N., Oka, T., Oseto, M., Hedlund, K.-O., Katayama, K., 2005. Intergenogroup recombination in sapoviruses. Emerg. Infect. Dis. 11, 1916–1920.

Hassan, F., Kanwar, N., Harrison, C.J., Halasa, N.B., Chappell, J.D., Englund, J.A., Klein, E.J., Weinberg, G.A., Szilagyi, P.G., Moffatt, M.E., Oberste, M.S., Nix, W.A., Rogers, S., Bowen, M.D., Vinje, J., Wikswo, M.E., Parashar, U.D., Payne, D.C., Selvarangan, R., 2019. Viral etiology of acute gastroenteritis in <2-year-old US children in the post-rotavirus vaccine era. J. Pediatric Infect. Dis. Soc. 8, 414–421.

Hoque, S.A., Nishimura, K., Thongprachum, A., Khamrin, P., Pham, N.T.K., Islam, M.T., Khandoker, N., Okitsu, S., Onda-Shimizu, Y., Dey, S.K., Maneekarn, N., Kobayashi, T., Hayakawa, S., Ushijima, H., 2022. An increasing trend of human sapovirus infection in Japan, 2009 to 2019: an emerging public health concern. J. Infect. Public Health 15, 315–320.

Irving, A.T., Ahn, M., Goh, G., Anderson, D.E., Wang, L.-F., 2021. Lessons from the host defences of bats, a unique viral reservoir. Nature 589, 363–370.

Katayama, K., Shirato-Horikoshi, H., Kojima, S., Kageyama, T., Oka, T., Hoshino, F.B., Fukushi, S., Shinohara, M., Uchida, K., Suzuki, Y., Gojobori, T., Takeda, N., 2002. Phylogenetic analysis of the complete genome of 18 Norwalk-like viruses. Virology 299, 225–239.

Katayama, K., Miyoshi, T., Uchino, K., Oka, T., Tanaka, T., Takeda, N., Hansman, G.S., 2004. Novel recombinant sapovirus. Emerg. Infect. Dis. 10, 1874–1876.

Katoh, K., Misawa, K., Kuma, K.-I., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Kroneman, A., Vega, E., Vennema, H., Vinje, J., White, P.A., Hansman, G., Green, K., Martella, V., Katayama, K., Koopmans, M., 2013. Proposal for a unified norovirus nomenclature and genotyping. Arch. Virol. 158, 2059–2068.

Kumar, S., Stecher, G., Li, M., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35, 1547–1549.

Kuroda, M., Masuda, T., Ito, M., Naoi, Y., Doan, Y.H., Haga, K., Tsuchiaka, S., Kishimoto, M., Sano, K., Omatsu, T., Katayama, Y., Oba, M., Aoki, H., Ichimaru, T., Sunaga, F., Mukono, I., Yamasato, H., Shirai, J., Katayama, K., Mizutani, T., Oka, T., Nagai, M., 2017. Genetic diversity and intergenogroup recombination events of sapoviruses detected from feces of pigs in Japan. Infect. Genet. Evol. 55, 209–217.

Le, S.Q., Gascuel, O., 2008. An improved general amino acid replacement matrix. Mol. Biol. Evol. 25, 1307–1320.

Li, J., Zhang, W., Cui, L., Shen, Q., Hua, X., 2018. Metagenomic identification, genetic characterization and genotyping of porcine sapoviruses. Infect. Genet. Evol. 62, 244–252.

Liu, B.L., Clarke, I.N., Caul, E.O., Lambden, P.R., 1995. Human enteric caliciviruses have a unique genome structure and are distinct from the Norwalk-like viruses. Arch. Virol. 140, 1345–1356.

Madeley, C.R., Cosgrove, B.P., 1976. Caliciviruses in man. Lancet 307, 199–200.

Nagai, M., Wang, Q., Oka, T., Saif, L.J., 2020. Porcine sapoviruses: pathogenesis, epidemiology, genetic diversity, and diagnosis. Virus Res. 286, 198025.

Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford University Press, Oxford, New York.

Oka, T., Wang, Q., Katayama, K., Saif, L.J., 2015. Comprehensive review of human sapoviruses. Clin. Microbiol. Rev. 28, 32–53.

Oka, T., Lu, Z., Phan, T., Delwart, E.L., Saif, L.J., Wang, Q., 2016. Genetic characterization and classification of human and animal sapoviruses. PloS One 11, e0156373.

Okitsu, S., Khamrin, P., Thongprachum, A., Hikita, T., Kumthip, K., Pham, N.T.K., Takanashi, S., Hayakawa, S., Maneekarn, N., Ushijima, H., 2021. Diversity of human sapovirus genotypes detected in Japanese pediatric patients with acute gastroenteritis, 2014–2017. J. Med. Virol. 93, 4865–4874.

Rambaut, A., 2018. FigTree. http://tree.bio.ed.ac.uk/software/figtree.

Simmonds, P., Karakasiliotis, I., Bailey, D., Chaudhry, Y., Evans, D.J., Goodfellow, I.G., 2008. Bioinformatic and functional analysis of RNA secondary structure elements among different genera of human and animal caliciviruses. Nucleic Acids Res. 36, 2530–2546.

Sunaga, F., Masuda, T., Aoki, H., Ito, M., Sano, K., Naoi, Y., Katayama, Y., Omatsu, T., Oba, M., Furuya, T., Shirai, J., Mizutani, T., Oka, T., Nagai, M., 2019. Complete genome sequencing and genetic characterization of porcine sapovirus genogroup (G) X and GXI: GVI, GVII, GX, and GXI sapoviruses share common genomic features and form a unique porcine SaV clade. Infect. Genet. Evol. 75, 103959.

Suzuki, Y., 2021. Effect of recombinations on changes in genotype proportions between norovirus seasons in Japan. Meta Gene 29, 100934.

Suzuki, Y., Kimura, H., Katayama, K., 2023. Shared clusters between phylogenetic trees for upstream and downstream regions of recombination hotspot in norovirus genomes. Gene Rep. 32, 101786.

Tohma, K., Kulka, M., Coughlan, S., Green, K.Y., Parra, G.I., 2020. Genomic analyses of human sapoviruses detected over a 40-year period reveal disparate patterns of evolution among genotypes and genome regions. Viruses 12, 516.

Wang, L., Marthaler, D., Fredrickson, R., Gauger, P.C., Zhang, J., Burrough, E.R., Petznick, T., Li, G., 2020. Genetically divergent porcine sapovirus identified in pigs, United States. Transbound. Emerg. Dis. 67, 18–28.

Wang, Q.H., Han, M.G., Funk, J.A., Bowman, G., Janies, D.A., Saif, L.J., 2005. Genetic diversity and recombination of porcine sapoviruses. J. Clin. Microbiol. 43, 5963–5972.

Yinda, C.K., Conceicao-Neto, N., Zeller, M., Heylen, E., Maes, P., Ghogomu, S.M., Van Ranst, M., Matthijnssens, J., 2017. Novel highly divergent sapoviruses detected by metagenomics analysis in straw-colored fruit bats in Cameroon. Emerg. Microbes Infect. 6, e38.

Yoshida, R., Nei, M., 2016. Efficiencies of the NJp, maximum likelihood, and Bayesian methods of phylogenetic construction for compositional and noncompositional genes. Mol. Biol. Evol. 33, 1618–1624.