Molecular Evolutionary Analyses of the RNA-Dependent RNA Polymerase (*RdRp*) Region and *VP1* Gene in Sapovirus GI.1 and GI.2

Fuminori Mizukoshi [1]*, Ryusuke Kimura [2,3], Tatsuya Shirai [1], Asumi Hirata-Saito [4], Eri Hiraishi [5], Kosuke Murakami [6], Yen Hai Doan [7], Hiroyuki Tsukagoshi [8], Nobuhiro Saruki [8], Takeshi Tsugawa [9], Kana Kidera [10], Yoshiyuki Suzuki [11], Naomi Sakon [12], Kazuhiko Katayama [10], Tsutomu Kageyama [7], Akihide Ryo [1], Hirokazu Kimura [3,5]*

[1] Department of Virology III, National Institute of Infectious Diseases, Musashimurayama-shi, Tokyo 208-0011, Japan; mzksh@niid.go.jp (F.M.); shirai@niid.go.jp (T.S.); aryo@niid.go.jp (A.R.)

[2] Department of Bacteriology, Graduate School of Medicine, Gunma University, Maebashi-shi 371-8511, Gunma, Japan; m2220015@gunma-u.ac.jp

[3] Advanced Medical Science Research Center, Gunma Paz University, Takasaki-shi 370-0006, Gunma, Japan; h-kimura@paz.ac.jp (H.K.)

[4] Department of Microbiology, Tochigi Prefectural Institute of Public Health and Environmental Science, Utsunomiya-shi, Tochigi 329-1196, Japan; saitoa01@pref.tochigi.lg.jp

[5] Department of Health Science, Gunma Paz University Graduate School of Health Sciences, Takasaki-shi 370-0006, Gunma, Japan; er.hiraishi@gmail.com

[6] Center for Emergency Preparedness and Response, National Institute of Infectious Diseases, Shinjuku-ku, Tokyo 162-8640, Japan; ko-mura@niid.go.jp

[7] Center for Emergency Preparedness and Response, National Institute of Infectious Diseases, Musashimurayama-shi, Tokyo 208-0011, Japan; yendoan@niid.go.jp (Y.H.D); tkage@niid.go.jp (T.K.)

[8] Gunma Prefectural Institute of Public Health and Environmental Sciences, Maebashi-shi 371-0052, Japan; tsuka-hiro@pref.gunma.lg.jp (H.T.); saruki-n@pref.gunma.lg.jp (N.S.)
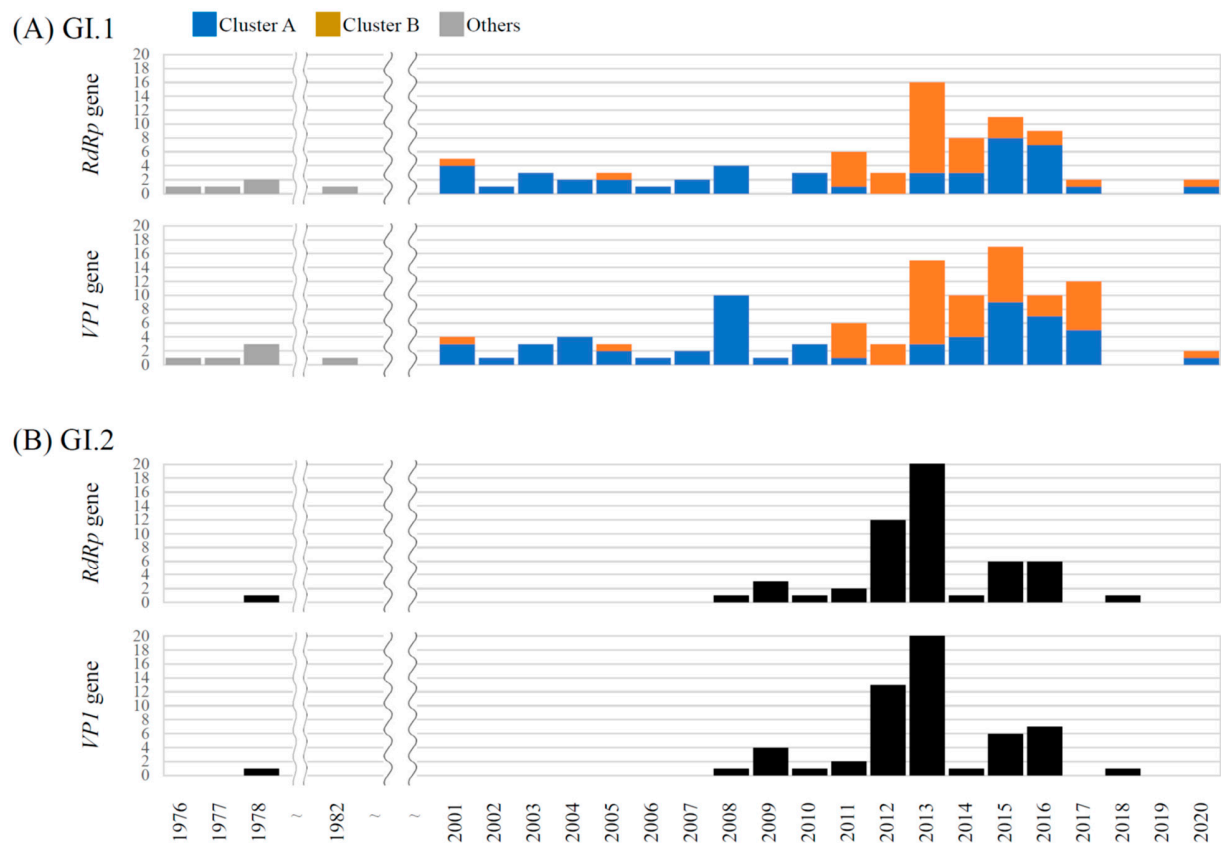
[9] Department of Pediatrics, Sapporo Medical University School of Medicine, Sapporo-shi 060-8543, Japan; tsugawat@sapmed.ac.jp

[10] Laboratory of Viral Infection Control, Ōmura Satoshi Memorial Institute, Graduate School of Infection Control Sciences, Kitasato University, 5-9-1, Shirogane, Minato-ku, Tokyo 108-8641, Japan; kidera.kana@st.kitasato-u.ac.jp (Kana Kidera); katayama@lisci.kitasato-u.ac.jp (Kazuhiko Katayama)

[11] Division of Biological Science, Department of Information and Basic Science, Graduate School of Sciences, Nagoya City University, Nagoya-shi, Aichi 467-8501, Japan; yossuzuk@nsc.nagoya-cu.ac.jp
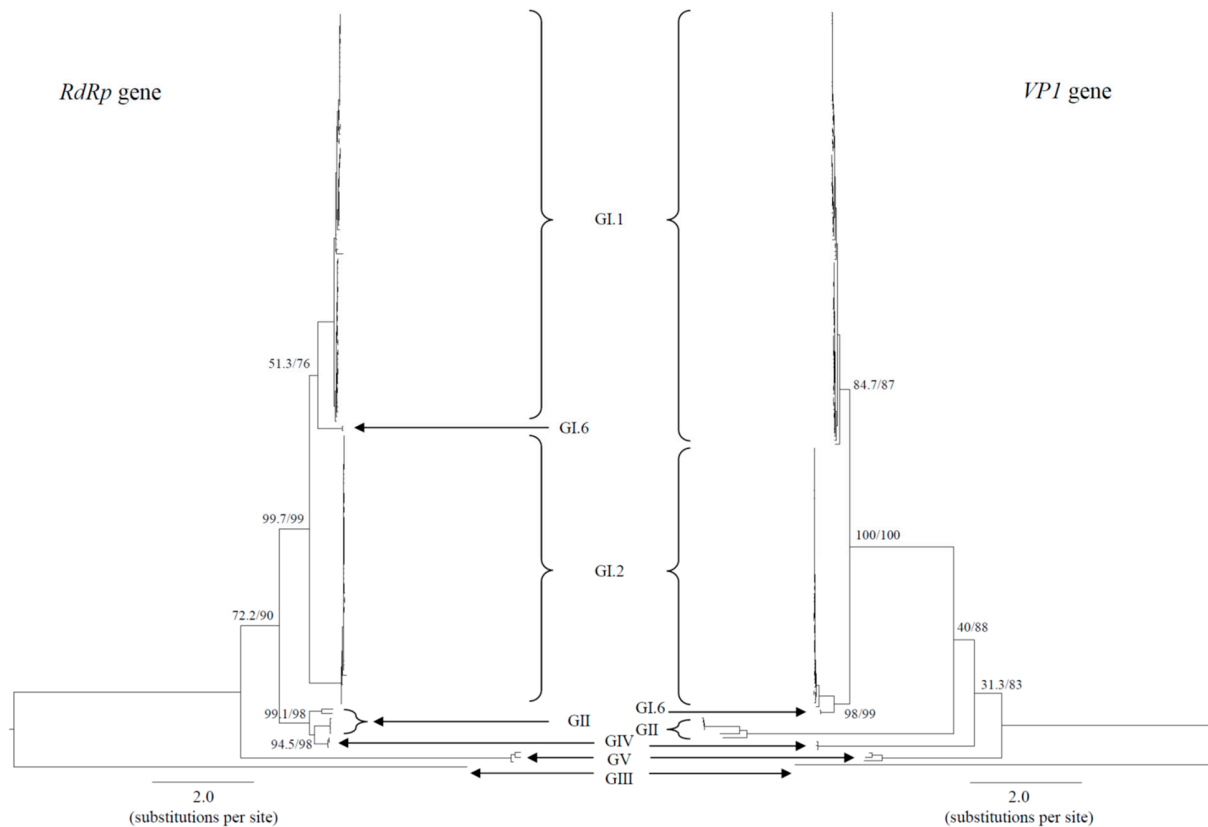
[12] Department of Microbiology, Osaka Institute of Public Health, Osaka 537-0025, Japan; nsakon@iph.osaka.jp

* Correspondence: mzksh@niid.go.jp; Tel.: +81-425-61-0771 (F.M.); h-kimura@paz.ac.jp; Tel.: +81-27-365-3366 (H.K.)

(A) GI.1

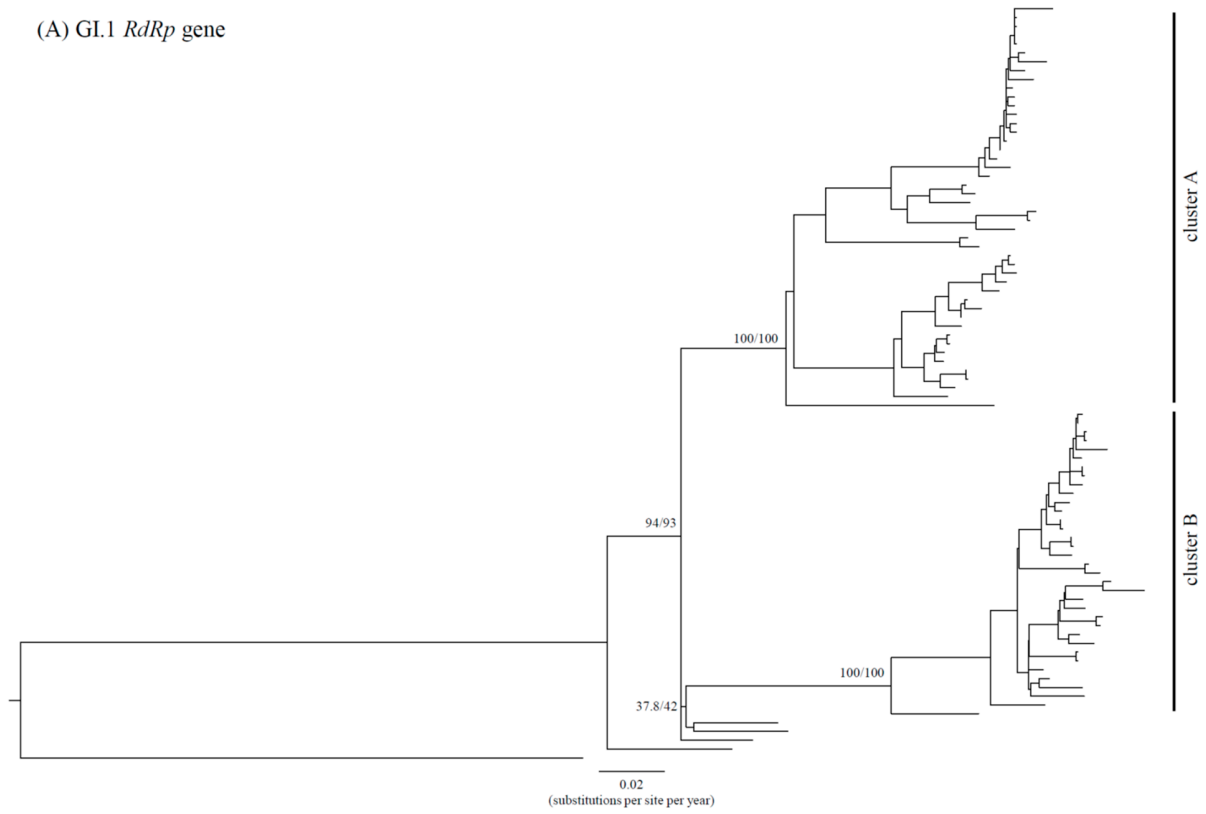Cluster A   Cluster B   Others

(B) GI.2

Supplementary Fig S1

Timeline distribution of sequences by collected year in this study. The *RdRp* region and *VP1* gene in HuSaV GI.1 (A) and GI.2 (B) genotypes were shown, respectively.
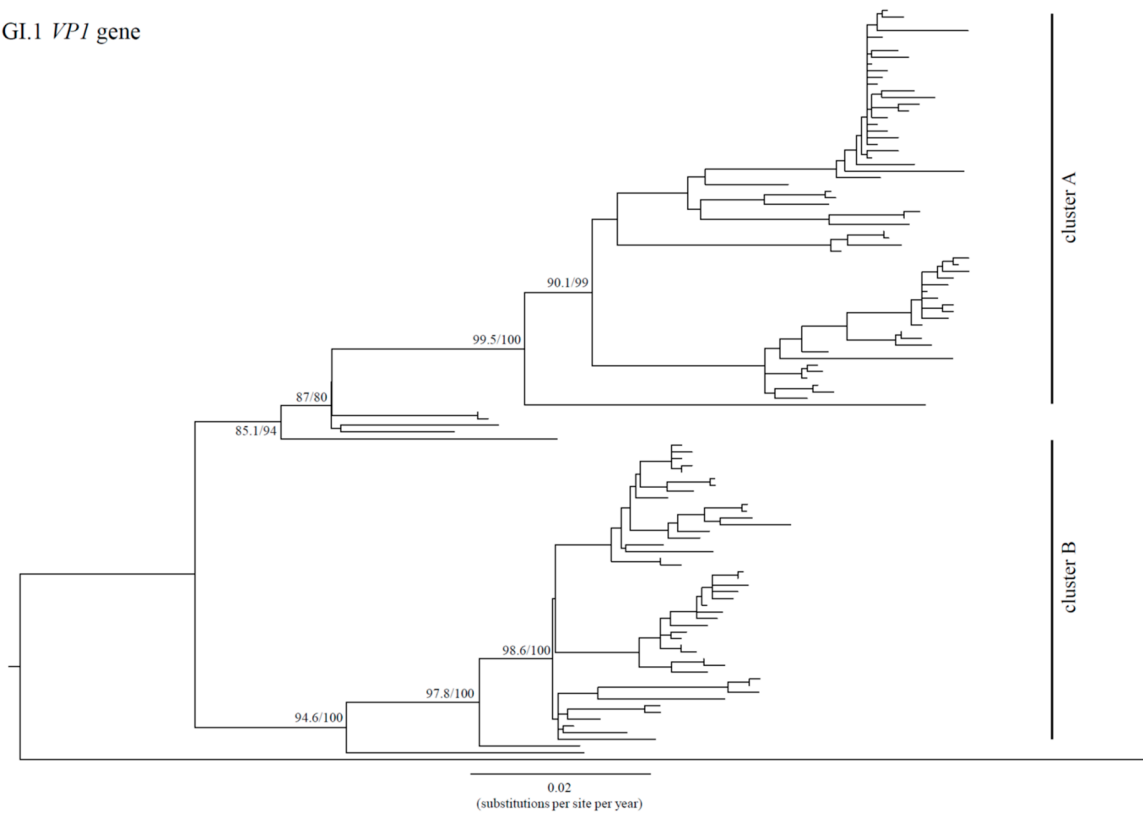
Supplementary Fig S2

Phylogenetic trees of the *RdRp* and *VP1* sequences of the various SaV genogroups (HuSaV GI.1 and GI.2, plus GI.6, GII, GIV, GV, and porcine SaV GIII) constructed by the ML method. Labels at the branch nodes show SH-aLRT / UltrafastBootstrap values. The scale bar shows nucleotide substitutions per site. The substitution model for the *RdRp* and the *VP1* gene was the TIM2+F+G4 model selected with ModelFinder.

3

(A) GI.1 *RdRp* gene
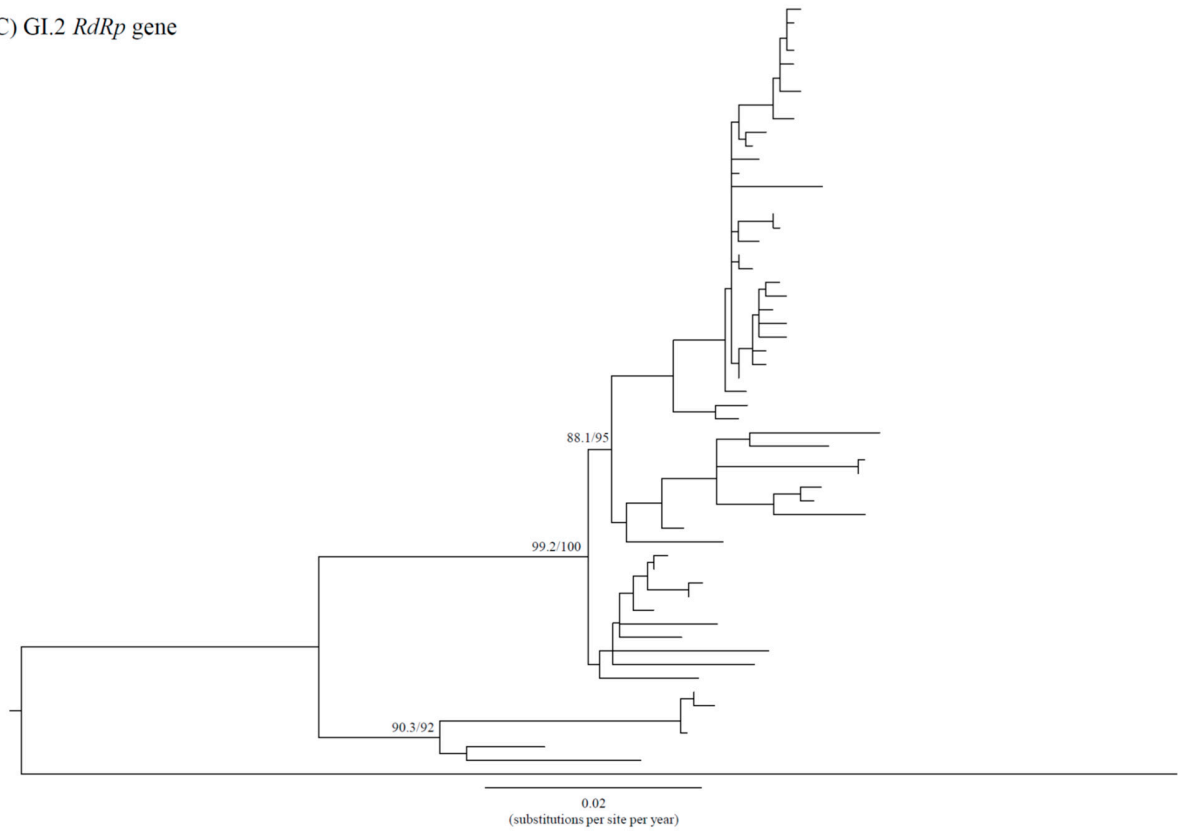
cluster A

100/100

94/93

37.8/42

100/100

cluster B

0.02
(substitutions per site per year)

(B) GI.1 *VP1* gene

cluster A

90.1/99

99.5/100

87/80

85.1/94

cluster B

98.6/100

97.8/100

94.6/100

0.02
(substitutions per site per year)

4

(C) GI.2 *RdRp* gene

88.1/95

99.2/100

90.3/92

0.02
(substitutions per site per year)

(D) GI.2 *VP1* gene

97.8/99

99.9/100

97.7/100

0.02
(substitutions per site per year)
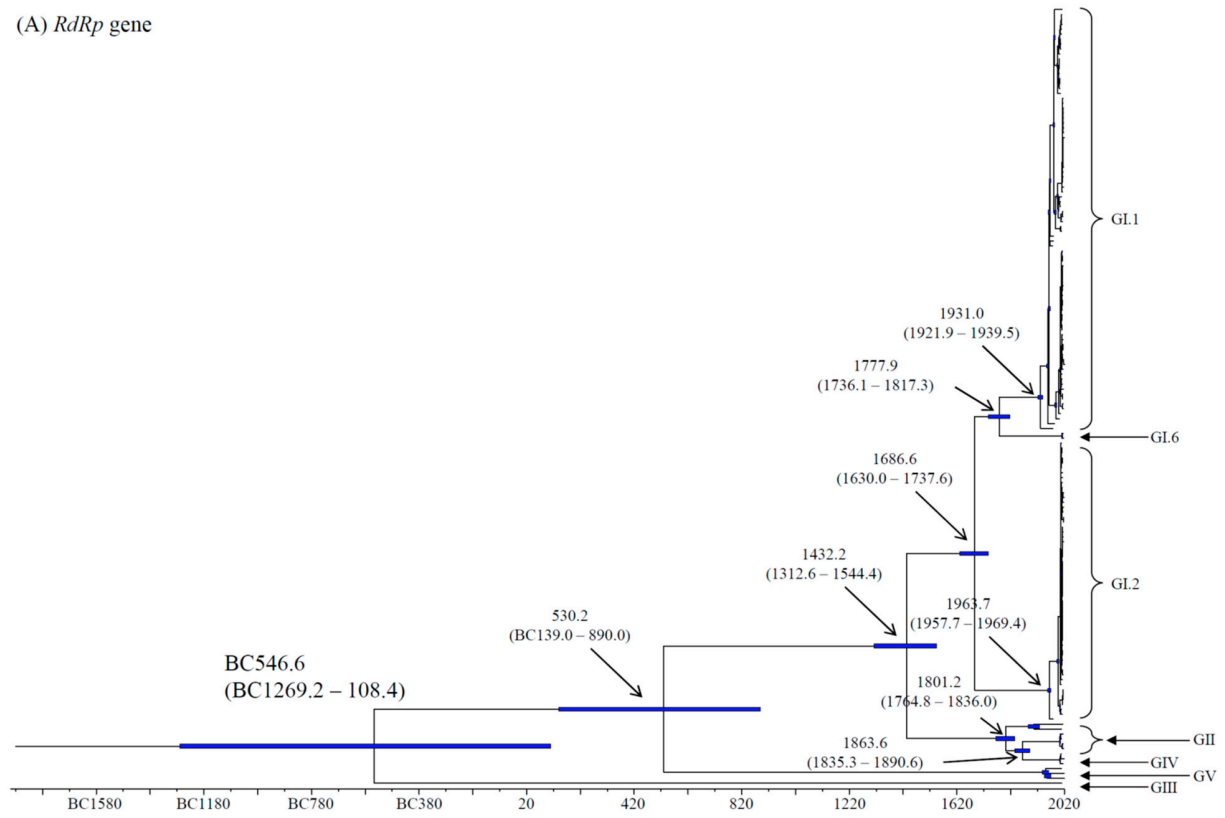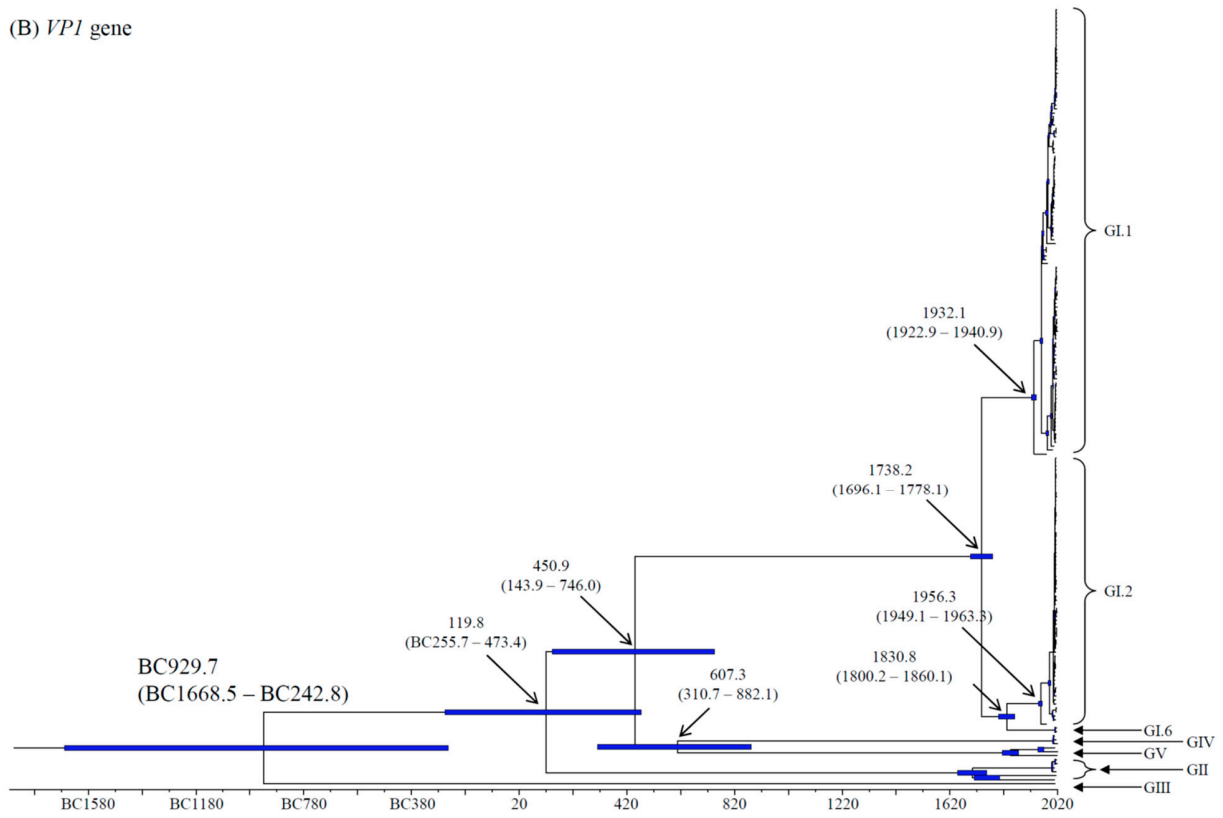
5

Supplementary Fig S3

Phylogenetic trees of *RdRp* regions (A) and *VP1* gene (B) of the HuSaV GI.1, and *RdRp* region (C) and *VP1* gene (D) of the HuSaV GI.2 constructed by the ML method. Labels at the branch nodes show SH-aLRT / UltrafastBootstrap values. The scale bar shows nucleotide substitutions per site. The substitution models selected with ModelFinder were SYM+I+R2, TIM3e+I+G4, TNe+I, and TIM2e+G4 for RdRp regions (A) and VP1 gene (B) of the HuSaV GI.1, and RdRp region (C) and VP1 gene (D) of the HuSaV GI.2, respectively.

(A) *RdRp* gene

1931.0
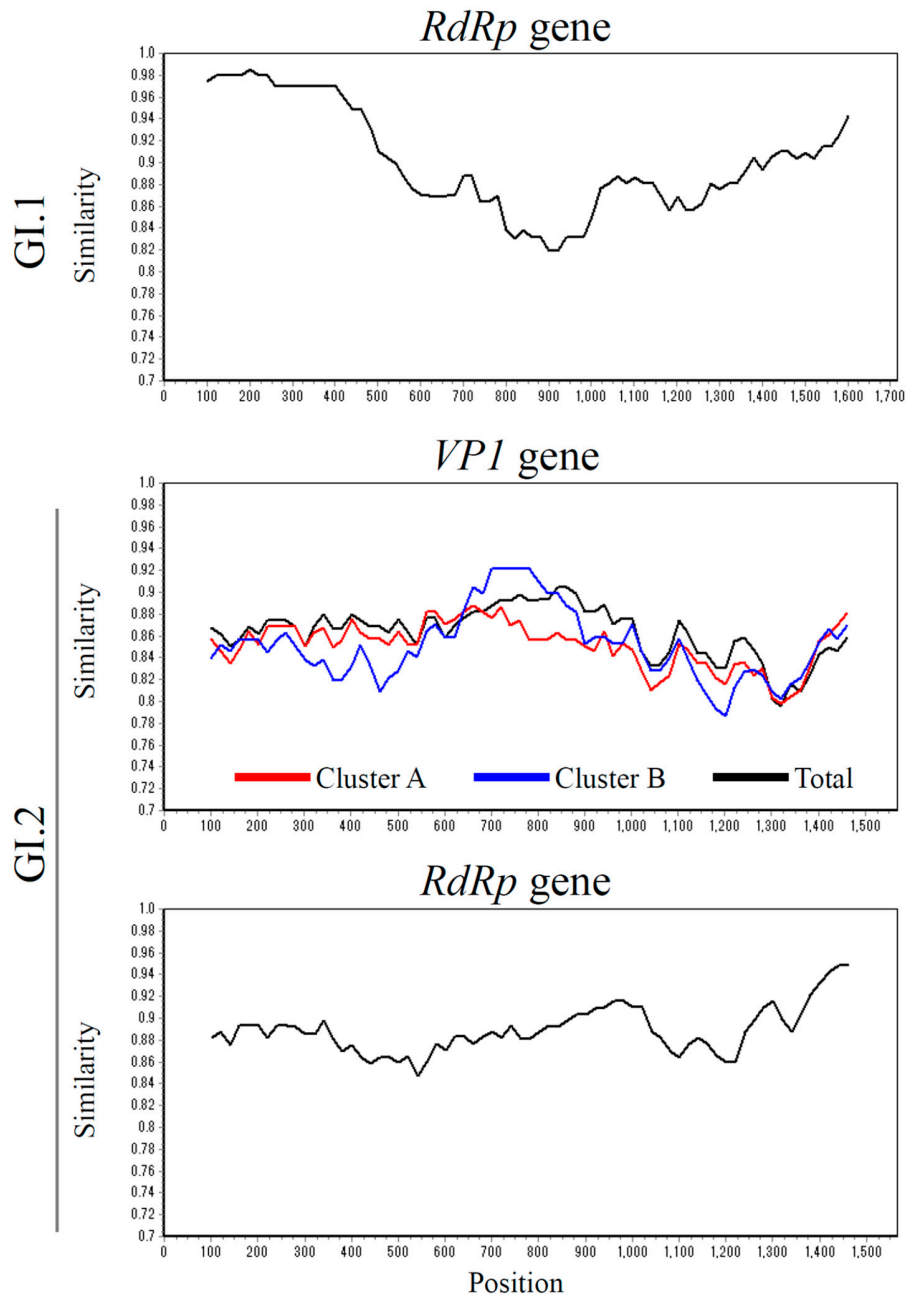(1921.9 – 1939.5)

1777.9
(1736.1 – 1817.3)

1686.6
(1630.0 – 1737.6)

1432.2
(1312.6 – 1544.4)

530.2
(BC139.0 – 890.0)

BC546.6
(BC1269.2 – 108.4)

1963.7
(1957.7 – 1969.4)

1801.2
(1764.8 – 1836.0)

1863.6
(1835.3 – 1890.6)

GI.1

GI.6

GI.2

GII
GIV
GV
GIII

BC1580   BC1180   BC780   BC380   20   420   820   1220   1620   2020

(B) *VP1* gene

1932.1
(1922.9 – 1940.9)

1738.2
(1696.1 – 1778.1)

450.9
(143.9 – 746.0)

119.8
(BC255.7 – 473.4)

BC929.7
(BC1668.5 – BC242.8)

607.3
(310.7 – 882.1)

1956.3
(1949.1 – 1963.3)

1830.8
(1800.2 – 1860.1)

GI.1

GI.2

GI.6
GIV
GV
GII
GIII

BC1580   BC1180   BC780   BC380   20   420   820   1220   1620   2020

7

Supplementary Fig S4

Time-scaled phylogenetic trees of *RdRp* region (A) and *VP1* gene (B) of t the various

SaV genogroups (HuSaV GI.1 and GI.2, plus GI.6, GII, GIV, GV, and porcine SaV GIII)

constructed by the Bayesian MCMC method. The horizontal axis represents time (years).

Blue bars indicate the 95% HPD for a branched year.

Supplementary Fig S5

Similarity plot analysis of the *RdRp* region of the HuSaV GI.1 (A), and *RdRp* region (B) and *VP1* genes (C) of the HuSaV GI.2. Nucleotide similarity to the prototype strain (GI.1: MN794208, collected in 1977; GI.2: MN794205, collected in 1978) was calculated using SimPlot analysis.