OPEN

Rank orders of mammalian pathogenicity-related PB2 mutations of avian influenza A viruses

Chung-Young Lee[1], Se-Hee An[1], Jun-Gu Choi (ID)[4], Youn-Jeong Lee[4], Jae-Hong Kim[1,3] & Hyuk-Joon Kwon[2,3,5*]

The PB2 gene is one of the key determinants for the mammalian adaptation of avian influenza A viruses (IAVs). Although mammalian pathogenicity-related mutations (MPMs) in PB2 genes were identified in different genetic backgrounds of avian IAVs, the relative effects of single or multiple mutations on viral fitness could not be directly compared. Furthermore, their mutational steps during mammalian adaptation had been unclear. In this study, we collectively compared the effects of individual and combined MPMs on viral fitness and determined their rank orders using a prototypic PB2 gene. Early acquired mutations may determine the function and potency of subsequent mutations and be important for recruiting multiple, competent combinations of MPMs. Higher mammalian pathogenicity was acquired with the greater accumulation of MPMs. Thus, the rank orders and the prototypic PB2 gene may be useful for predicting the present and future risks of PB2 genes of avian and mammalian IAVs.

Waterfowl are reservoirs for influenza A viruses (IAVs), and close interaction between waterfowl and other animals causes occasional cross species transmission to result in successful settle-down by acquiring host adaptive mutations in their eight segmented genomes (PB2, PB1, PA, HA, NP, NA, M, and NS)[1]. In particular, the PB2 protein, which is involved in cap snatching of the host mRNA, has been regarded as one of the key molecules to overcome species-specific host barriers[2–4]. PB2 interacts with various host proteins, including importin-α for nuclear localization and ANP32A for polymerase stabilization and mitochondrial antiviral signaling proteins (MAVS), Tu elongation factor, and mitochondrial (TUFM) for anti-viral activity[5–8].

To date, many mammalian pathogenicity-related mutations (MPMs, D9N, I147T, E158G, E192K, A199S, D253N, T271A, K339T, F404L, K526R, A588I, A588T, G590S, Q591R, Q591K, E627K, A674T, D701N, K702R, and S714R, etc.) in PB2 have been reported in different domains of the PB2 protein[3,4,9–27]. The E627K mutation most potently increases polymerase activity, replication efficiency and pathogenicity of avian IAVs in mammalian hosts[3,4,17]. However, avian IAVs cumulatively have acquired multiple mutations other than E627K, and there may be additional requirements to achieve successful adaptation to mammals. As MPMs were identified and verified in different genetic contexts, the relative effects of MPMs on viral fitness needed to be compared in the same genetic backgrounds to omit epistasis[28]. In addition, swine-origin pandemic H1N1 viruses in 2009 [A(H1N1) pdm2009] showed human pathogenicity without E627K mutation, and different mutational steps to acquire mammalian pathogenicity were proposed[21]. Human and swine IAVs possess already multiple MPMs, but their effects on mammalian pathogenicity have never been compared collectively. Furthermore, the mutational steps of MPMs during mammalian adaptation were unclear.

Previously, we found a prototypic avian PB2 gene (01310 PB2) that completely attenuated PR8-derived recombinant virus (7 + 1 reassortant) in mice and did not possess any MPMs except 661A and 683T, and identified the

[1]Laboratory of Avian Diseases, College of Veterinary Medicine, Seoul National University, 08826, Seoul, Republic of Korea. [2]Department of Farm Animal Medicine, College of Veterinary Medicine, Seoul National University, 08826, Seoul, Republic of Korea. [3]Research Institute for Veterinary Science, College of Veterinary Medicine, Seoul National University, 08826, Seoul, Republic of Korea. [4]Avian Disease Division, Animal and Plant Quarantine Agency, 177, Hyeoksin 8-ro, Gyeongsangbuk-do, 39660, Republic of Korea. [5]Farm Animal Clinical Training and Research Center (FACTRC), GBST, Seoul National University, Kangwon-do, Republic of Korea. *email: kwonhj01@snu.ac.kr

Category	No. of compared sequence	Domain/% of mammalian pathogenic mutation of PB2 gene																			
		N1	N2	Lid			Mid		Cap-binding		Cap-627 linker	627					NLS				
		D9N	I147T	E158G	E192K	A199S	D253N	T271A	K339T	E404L	K526R	T588I	A588T	G590S	Q591R	Q591K	E627K	A674T	D701N	K702R	S714R
Bird	11,649	0.75	11.22	0.02	0.04	0.12	0.04	0.40	8.80	0.18	0.09	0.09	4.02	8.40	0.15	0.15	2.77	0.41	0.10	3.75	0.01
Bird-Human	293	0.00	16.72	0.00	1.37	2.39	0.00	0.34	34.13	0.00	20.48	0.34	5.46	7.85	0.34	2.39	40.96	0.34	6.14	1.71	0.00
Pig	3,191	3.35	79.16	0.06	0.03	1.54	0.03	79.66	0.56	0.06	3.17	4.86	74.55	77.28	79.35	0.97	1.25	1.69	18.02	2.16	0.00
pdmH1N1	4,721	0.32	100	0.02	0.02	0.00	0.02	99.94	0.00	0.00	2.37	2.50	97.12	99.49	99.92	0.00	0.00	0.06	0.06	0.04	0.04
Human	7,748	95.43	0.49	0.00	0.03	99.69	0.00	99.72	0.04	0.00	94.72	77.58	21.05	74.21	0.31	0.01	99.63	97.44	0.25	98.94	0.08

Figure 1. Mammalian pathogenic PB2 mutations and frequency in different hosts. The location and frequency of 20 MPMs of 18 amino acid residues of PB2 were compared and the frequency difference was visualized by heat-map, 0 (yellow) to 100% (red).

minimum essential mutations (I66M, I109V, and I133V, MVV mutations) to acquire additional MPMs such as Q591R/K, E627K, and D701N[29,30]. The minimum essential mutations increased viral replication efficiency in both avian and mammalian cells and were postulated to increase the trimeric integrity of polymerase complex[29]. As epistatic mutations blur the effects of certain mutations, the use of the prototypic PB2 gene may be crucial to minimize epistasis[28].

Therefore, in this study, we grafted each of 20 MPMs of 18 amino acid residues related to mammalian pathogenicity in the 01310PB2 gene and compared their effects on polymerase activity. Then, we generated PR8-derived recombinant viruses with selected mutations in 01310PB2 to compare replication efficiency in mammalian hosts. According to the modified median-joining network analysis based on profiles of the twenty MPMs and chronological data, we hypothesized mutational steps of MPMs in major PB2 genotypes of A(H1N1) pdm2009 and seasonal flu viruses. We compared polymerase activities of the hypothetical intermediate and final combinations of MPMs acquired during adaptation to pigs and humans and generated PR8-derived recombinant viruses possessing major combinations of MPMs to compare their effects on viral replication efficiency and pathogenicity in mammalian hosts.

Results

The frequency of each MPM in different categories of PB2 genes.
We collected 27,602 PB2 genes possessing MVV mutations that comprised 84.7% of PB2 genes (32,591) in the Influenza Research Database (IRD)[29]. We classified them into 5 categories, Bird (avian IAVs), Bird-Human [avian IAVs isolated form humans (H5, H6, H7, H9, and H10 IAVs)], Pig (swine IAVs), pdmH1N1 (H1N1 IAVs during 2009 pandemic), and Human (human IAVs), and calculated the frequencies of their MPMs. The twenty mutations of 18 amino acid residues, D9N, I147T, E158G, E192K, A199S, D253N, T271A, K339T, F404L, K526R, A588I/T, G590S, Q591R/K, E627K, A674T, D701N, K702R, and S714R, were located in different domains of PB2 and the frequency difference was visualized by heat map (Fig. 1).

The frequency of each mutation was apparently different between different categories. The frequencies of the mutations D9N, A199S, K526R, A588I, E627K, A674T, and K702R were highest in Human relative to the other categories, but the frequencies of mutations I147T, A588T, and Q591R were higher in the Pig and pdmH1N1 categories than in the others. Although the frequency of D701N in Pig (18.02%) was not high, it was relatively higher than that in other categories (0.06–0.25%). Only T271A and G590S were commonly frequent in the Pig, pdmH1N1 and Human categories. Although most of mutations were less frequent in Bird than in other categories, I147T, K339T, A588T, G590S, E627K, and K702R were relatively more frequent than the other mutations. In the Bird-Human category, the frequencies of E627K (from 2.77 to 40.96%), K339T (from 8.8 to 34.13%), and K526R (from 0.09 to 20.48%) increased steeply in comparison with the corresponding mutations in the Bird category. The frequencies of E158G, E192K, D253N, F404L, Q591K, and S714R were very low, less than 2.4%, in all categories.

The effect of each MPM on polymerase activity, replication efficiency and pathogenicity in mammalian hosts.
To compare the effect of each PB2 mutation on polymerase activity, we selected 19 MPMs except A674T, which had no effect on polymerase activity[29]. We measured the polymerase activity by using an *in vitro* mini-genome assay in 293T cells (Fig. 2A). Except for the F404L, A588T, and S714R mutations, all PB2 mutations significantly increased polymerase activity relative to the wild-type 01310 PB2 gene. In particular, the E627K, E158G, Q591K, and D253N mutations showed the highest polymerase activities in order and increased polymerase activity by more than 20-fold in comparison with that of the 01310 PB2 gene (Fig. 2A). D701N, E192K and K526R showed approximately 10-fold increases in polymerase activities, followed by A588I and Q591R. Although D9N, I147T, A199S, T271A, K339T, G590S, and K702R showed significantly higher polymerase activity than wild-type 01310 PB2, they increased polymerase activities by 3-fold or less. The rank order of the effects of PB2 mutations on polymerase activity was summarized in Table S2.

To investigate the effects of the MPMs on viral replication, we selected 10 mutations that increased the polymerase activity by more than 3-fold (E158G, E192K, D253N, T271A, K526R, T588I, Q591K, Q591R, E627K, and D701N) and generated PR8-derived recombinant viruses containing the 10 mutated 01310 PB2 genes. In
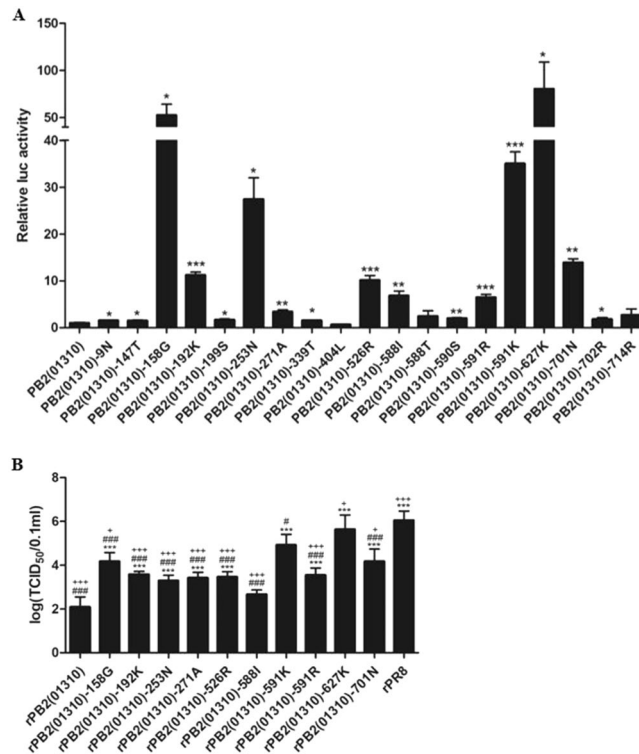
Figure 2. Effect of single MPMs on polymerase activity and replication efficiency in MDCK cells. Polymerase activity was measured using mini-genome assays in 293T cells (**A**). The data were normalized to the polymerase activity of the wild-type 01310 PB2 gene. Statistical significance was calculated using one-way ANOVA ($^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$). Replication efficiency of PR8-derived recombinant viruses possessing single MPMs in 01310PB2 gene in MDCK cells at 37 °C (**B**). Wild-type and mutant viruses (10^7 $EID_{50}/0.1$ mL) were used to infect MDCK cells, and the $TCID_{50}$ was measured at 5dpi. Statistical significance was analyzed by one-way analysis of variance with Bonferroni correction (compared to rPB2(01310), $^{***}P < 0.001$; compared to rPB2(01310)-627K, $^{\#}P < 0.05$, $^{\#\#\#}P < 0.001$; compared to rPB2(01310)-591K, $^{+}P < 0.05$, $^{+++}P < 0.001$). The data are the average of triplicate data $\pm$ s.d. from one experiment.

MDCK cells, most of the mutant viruses showed significantly higher replication efficiency than the wild-type virus, rPB2(01310), except rPB2(01310)-588I (P < 0.05) (Fig. 2B). The E627K mutation increased virus titers significantly in comparison with other mutations. The Q591K mutation more greatly increased virus titer than the E158G and D701N mutations (Fig. 2B). The rank order of the effects of MPMs on replication efficiency in MDCK cells was summarized in Table S2.

We investigated the effect of the 10 PB2 mutations on mouse pathogenicity by observing body-weight loss of BALB/c mice (Fig. 3). Most of the single mutations introduced into the prototypic 01310 PB2 gene did not cause body-weight loss during the observation period, but rPB2(01310)-E627K slightly induced the body weight loss of mice up to 12% (Fig. 3A). Only rPR8 caused mortality during the observation period (Fig. 3B). However, all of the single amino acid mutant viruses could replicate in the lungs on 3 and/or 6 dpi in contrast to the parent virus, rPB2(01310) (Table 1). As expected, rPB2(01310)-627K showed the highest virus titers among other viruses on 3 dpi, but rPB2(01310)-701N reached to the highest virus titers on 6 dpi. The effects of E158G, 192K, 271A and Q591K on virus replication in the lungs were less apparent than those of E627K and D701N but higher than those of D253N, A588I and Q591R. The rank order of the effects of MPMs on virus replication in the lungs was summarized in Table S2.

Inference of the evolutionary steps of swine and human PB2 genes. To understand the mutational steps of MPMs, we performed median joining network analysis using the profiles of 20 mutations of PB2 genes. According to the results, predominant PB2 genotypes in avian, A(H1N1) pdm2009, and human seasonal flu viruses were apparent: MVV, SIB-5–1 and HIB-9-1, respectively. The avian MVV genotype may have evolved to become SIB-5-1 and HIB-9-1 through multiple mutational steps (Fig. 4). However, the chronological data of the PB2 genes were not reflected in the median joining network analysis tree. Therefore, we modified the order of mutational steps from MVV to SIB-5-1 and HIB-9-1 based on chronology-frequency consideration of PB2 mutations as described below.

In the case of the SIB-5-1 genotype, avian IAVs possessing MVV-G590S, MVV-I147T, and MVV-G590S-I147T were first observed in 1977 [A/canvasback_duck/ALB/274/1977 (H4N6), CY004728)], 1968 [A/turkey/Wisconsin/1/1968 (H5N9), CY080514] and 1986 [A/ blue_winged_teal/LA/B228/1986 (H1N1), EU743313], respectively. Therefore, early PB2 genes bearing MVV-G590S and MVV-I147T have coexisted, and
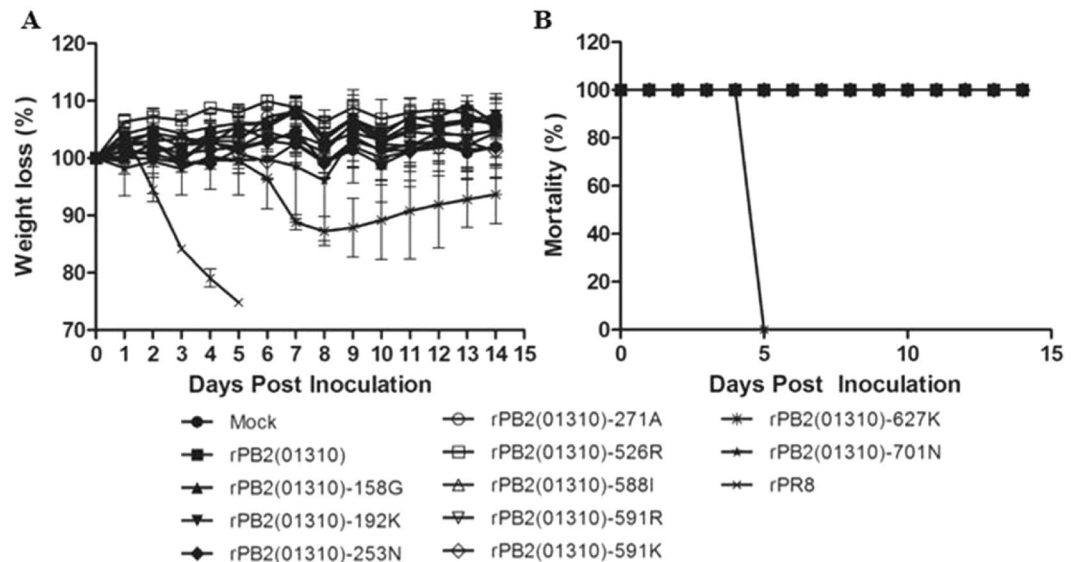
Figure 3. Effects of single MPMs on mouse pathogenicity. The pathogenicity of single amino acid mutant viruses was observed by body weight loss (**A**) and mortality (**B**) of infected mice. Five 6-week-old BALB/c mice were challenged with 1.0×10^6 EID$_{50}$ of each virus or PBS (mock). Weight loss was observed for 14 days, and mice showing more than 20% body weight loss were regarded as dead. The average weight loss $\pm$ s.d. was measured by comparing to the initial weight of each mouse.

Virus	3 dpi		6 dpi	
	Positive rate*	Virus titer (logEID$_{50}$/0.1 mL)	Positive rate	Virus titer (logEID$_{50}$/0.1 mL)
Mock (PBS)	0/3	<0.5†	0/3	<0.5
rPB2(01310)	0/3	<0.5	0/3	<0.5
rPB2(01310)-158G	3/3	3.2 ± 1.7	3/3	3.1 ± 0.9
rPB2(01310)-192K	3/3	2.5 ± 1.3	3/3	2.8 ± 2.1
rPB2(01310)-253N	0/3	<0.5	1/3	0.8 ± 0.6
rPB2(01310)-271A	3/3	2.4 ± 0.5	3/3	3.5 ± 1.5
rPB2(01310)-526R	2/3	1.3 ± 1.4	2/3	2.3 ± 1.5
rPB2(01310)-588I	1/3	1.1 ± 1.0	1/3	1.5 ± 1.7
rPB2(01310)-591R	2/3	<0.5	3/3	1.5 ± 1.1
rPB2(01310)-591K	3/3	2.9 ± 1.2	3/3	3.4 ± 0.4
rPB2(01310)-627K	3/3	5.0 ± 0.4	3/3	4.4 ± 0.4
rPB2(01310)-701N	3/3	3.8 ± 0.9	3/3	4.8 ± 0.3
rPR8	3/3	6.1 ± 0.5	1/1	5.0‡

Table 1. Comparison of the viral replication efficiency of 01310 PB2 single amino acid mutant viruses in mouse lungs. *Number of positive samples/number of mice (virus titer of the pooled lung tissues, log10 EID$_{50}$/0.1 mL). †Below the limit of assay detection. ‡Data from only one mouse due to the death of two mice.

MVV-G590S-I147T may have appeared after acquiring the I147T or G590S mutation (Fig. 4). Considering the relatively low frequency and absence of MVV-G590S-I147T among avian and mammalian IAVs, respectively (relatively small size and yellow circle), it may have acquired additional Q591R, T271A and A588T successively due to the very low frequencies of the intermediates (MVV-G590S-I147T-Q591R, MVV-G590S-I147T-Q591R-T271A and MVV-G590S-I147T-Q591R-T271A-A588T) of MVV-G590S-I147T and SIB-5-1 (Fig. 4). Most A(H1N1) pdm2009 viruses possessed both Q591R and T271A, but only two strains possessed either Q591R or T271A. Therefore, MVV-G590S-I147T may have acquired Q591R and T271A mutations nearly simultaneously. After the following acquisition of A588T, SIB-5-1 may have become predominant, and evolved to become SIB-5-2 or SIB-6 by acquisition of T588I or K526R. Therefore, PB2 genes of A(H1N1) pdm2009 viruses may have acquired G590S/I147T, Q591R/T271A, A588T, T588I, and K526R mutations cumulatively (Fig. 4).

In the case of HIB-9-1 the PB2 gene of A(H1N1) pdm1918 already possessed MVV, A199S, E627K and K702R, and we inferred firstly the cumulative order of the A199S, E627K and K702R mutations (Table S3). The E627K mutation may have been acquired earlier due to an apparent higher frequency than that of A199S (Fig. 1). In the case of swine PB2 genes with TIV instead of MVV, E627K-A199S were observed without K702R and D9N, and A199S may precede K702R and D9N (Table S4). However, whether K702R or D9N was acquired next to
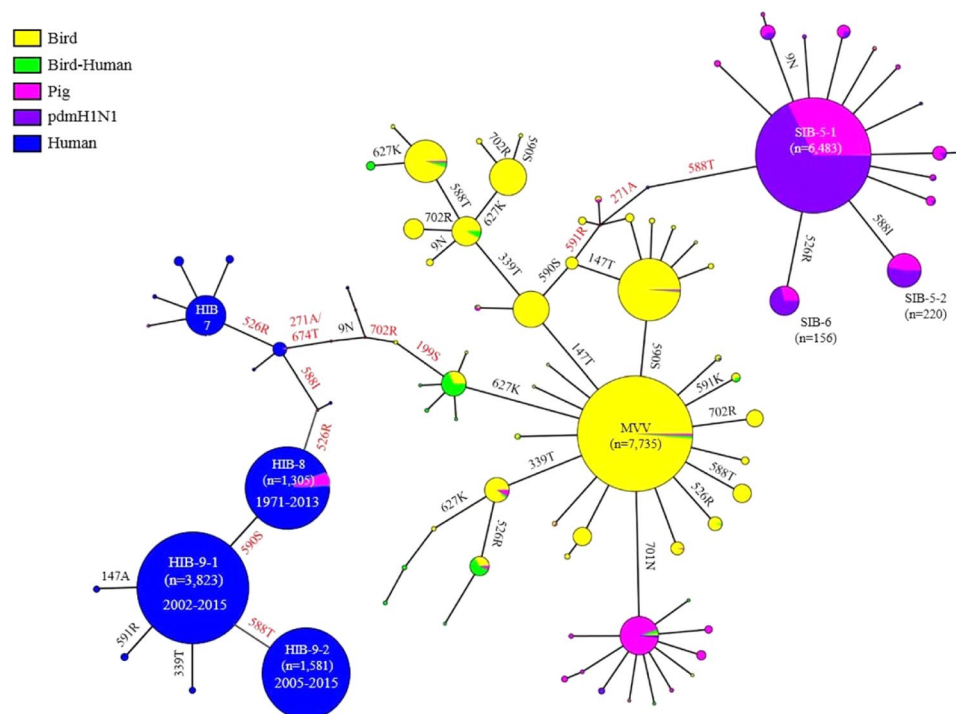
Figure 4. Phylogenetic network analysis of avian, swine and human PB2 genes and mutational steps of MPMs during mammalian adaptation. The profiles of MPMs of PB2 genes (n = 24,825) were analyzed by median joining network analysis with Network 5.0.0. Major genotypes were represented in the circles, and the size of circle is proportional to its frequency. Five different categories were represented in different colors. According to the manual chronology and frequency analysis of MPMs, mutational steps of major genotypes (SIB-5-1, SIB-5-2, SIB-6, HIB-8, HIB-9-1, and HIB-9-2) were hypothesized and represented in red color letters.

A199S was hard to determine because swine PB2 genes with TVV instead of MVV during the 1930s had K702R without D9N, while in contrast, swine TIV PB2 in the 1970s only had D9N without K702R (Table S5). However, A/Brevig Mission/1/1918 (H1N1) possesses only K702R without D9N, and we represented K702R as earlier than D9N (Table S1, Fig. 4). Out of the second earliest human H1N1 viruses during 1933-1936 (8 strains), five strains possessed K702R without T271A, and two strains have both T271A and K702R. Only the PR8 strain (4 out of 5 registered PR8 sequences in the database) possesses T271A without K702R. Therefore, K702R may precede T271A. In addition, five strains during 1933-1936 possess A674T without T271A, with 3 strains possessing both. Therefore, A674T may precede T271A (Table S3). The PB2 of A(H3N2) pdm1968 viruses possessed both T271A and A588I mutations, and it could not be determined which one was acquired earlier. Additionally, single point mutations do not cause A to I but do cause V to I mutations at 588 residue. Therefore, higher frequency of 588V than 588A in the incomplete mutant PB2 may support V588I mutation (Table S6). The HIB-8 genotype might have acquired K526R after the T271A and V588I mutations (Fig. 4, Table S7). The HIB-8 genotype was prevalent among seasonal flu viruses during 1971-2013, but the HIB-9-1 and HIB-9-2 genotypes were prevalent during 2002-2015 and 2005-2015, respectively. HIB-9-1 might have evolved from HIB-8 by acquisition of G590S and HIB-9-2 from HIB-9-1 by acquisition of the I588T mutation (Fig. 4). HIB-9-1 and HIB-9-2 still evolved to acquire additional MPMs such as I147T, K339T, Q591R or D701N (Table S8). Therefore, MVV PB2 genes may have acquired E627K, A199S, K702R, D9N, A674T, T271A/V588I, K526R, G590S, and I588T cumulatively (Fig. 4, Table S1). Although the predominant SIB- and HIB-series of PB2 genotypes started from different early mutations (MVV-G590S/I147T and MVV-E627K), the last half of the accumulation steps of SIB- and HIB-series genotypes shared similar mutations: T271A, A/V588T/I, and K526R. In addition, HIB-9-1 genotype PB2 genes have evolved by acquiring additional mutations such as I147T, Q591R, or K339T (Fig. 4).

The effects of combined multiple MPMs on polymerase activity. Previously, we found that the introduction of MVV mutations to prototypic 01310PB2 led to the acquisition of additional mutations such as Q591K, Q591R, E627K, and D701N in the lungs of BALB/c mice[29]. Moreover, the Q591K, Q591R, E627K, and D701N mutations are key factors for AIVs to overcome the host barrier[20–22,31–35]. Therefore, we compared the collaborative effects of MVV mutation with each Q591K, Q591R, E627K, and D701N mutation on polymerase activity. In combination with the MVV mutation, Q591K, Q591R, E627K, and D701N each showed increased polymerase activity at 37 °C, and MVV mutation had a synergistic relationship with each tested mutation (Fig. S1). The polymerase activity of MVV-E627K was highest, followed by MVV-Q591K, MVV-D701N, and MVV-Q591R. In addition, the single mutation of I147T (SIB-1-3) and A199S (HIB-1-3) significantly increased polymerase activity only at 37 °C, but T271A (SIB-1-1) increased polymerase activity at both 33 °C and 37 °C

(Fig. 5). The polymerase activity of I147T (SIB-1-3) and A199S (HIB-1-3) was similar to each other, but the polymerase activity of T271A (SIB-1-1) was higher than that of I147T (SIB-1-3) and A199S (HIB-1-3) and lower than that of Q591R (SIB-1-2).

The synergistic effect of T271A and Q591R (SIB-2-1) apparently caused a significant increase in polymerase activity at 37 °C in comparison with two other pairs, T271A-I147T (SIB-2-2) and Q591R-I147T (SIB-2-3). Although SIB-3 (I147T), SIB-4 (I147T-A590S) and SIB-5 (I147T-A590S-A588T) possessed additional mutations, their polymerase activities did not show significant differences from that of SIB-2-1. However, acquisition of the A588I and K526R mutations by SIB-5-2 and SIB-6 significantly increased their polymerase activity at 33 °C in comparison with that of SIB-5-1.

The double mutations of the HIB-2-3 genotype, A271T and A199S, increased polymerase activity at 37 °C by approximately 2-fold in comparison with the HIB-1-3 and HIB-1-2 genotypes. The single E627K mutation of the HIB-2-3 genotype led to significantly higher polymerase activity than HIB-1-1 with double mutations (A199S and T271A). Additional A199S and/or T271A mutations to E627K in the HIB-2-2, HIB-2-1 and HIB-3 genotypes increased polymerase activity in comparison with that of the HIB-2-3 genotype, and these findings suggest the 199S and 271A mutations have synergistic cooperation with the E627K mutation. However, there was no significant difference in polymerase activity among HIB-3, 4, 5, 6, 7, 8, 9-1, and 9-2 at any temperature. Therefore, the effects of additional mutations such as K702R, A674T, D9N, K526R, V588I I588T, and G590S on polymerase activity were indistinguishable.

The effects of the predominant combinations of MPMs on replication efficiency in mammalian hosts and mouse pathogenicity.

PR8-derived recombinant viruses containing the predominant PB2 genotypes, rPB2(01310)-SIB-5-1, rPB2(01310)-SIB-5-2, rPB2(01310)-SIB-6, rPB2(01310)-HIB-8, rPB2(01310)-HIB-9-1, and rPB2(01310)-HIB-9-2, were generated, and additionally their nucleotide sequences of PB1 and PA genes were verified to exclude the presence of co-evolutionary mutations. We compared their growth kinetics in A549 cells. rPB2(01310)-SIB-5-2 and rPB2(01310)-SIB-6 showed significantly higher replication efficiency than rPB2(01310)-SIB-5-1 at 24, 36, and 48 hpi in A549 cells (Fig. 6, $P < 0.05$). Therefore, the K526R and A588I mutations were involved in more efficient replication in A549 cells. However, rPB2(01310)-HIB-8, rPB2(01310)-HIB-9-1, and rPB2(01310)-HIB-9-2 showed indistinguishable replication efficiencies over all time intervals.

To compare the mouse pathogenicity of the predominant PB2 genotypes, we measured the MLD_{50} by intranasal inoculation of the recombinant viruses to BALB/c mice. The MLD_{50} values of rPB2(01310)-SIB-5-1, rPB2(01310)-SIB-5-2 and rPB2(01310)-SIB-6 were $>10^5$, 1.0×10^4 and 5.6×10^3 EID_{50}, respectively. rPB2(01310)-SIB-5-2 and rPB2(01310)-SIB-6 were more pathogenic than rPB2(01310)-SIB-5-1, and rPB2(01310)-SIB-6 was comparable to rPR8 (Fig. 7). The MLD_{50} values of rPB2(01310)-HIB-8, rPB2(01310)-HIB-9-1 and rPB2(01310)-HIB-9-2 were 5.6×10^2, 1.0×10^3 and 3.2×10^2 EID_{50}, respectively. PR8-derived viruses possessing the HIB-8, 9-1 and 9-2 PB2 genotypes were more pathogenic than SIB-series and rPR8 viruses.

Discussion

The mutation rate per site may be similar, but the effect of each mutation on viral fitness may be different [28,36]. Mutations that increase viral fitness may become prevalent among competing mutations. During several pandemics and continuous endemics, frequent infection and transmission among humans may facilitate competition between IAVs and impose selective pressures for more competent PB2 mutations. Therefore, more than 94% frequencies of D9N, A199S, T271A, K526R, E627K, A674T, and K702R mutations in the Human category may reflect the results of intensive competition and selection of such mutations.

The earliest swine IAVs date back to 1918 and are believed to be closely related to pandemic and endemic viruses[37]. However, 'triple reassortant' H3N2 viruses probably appeared via reassortment of the HA, NA and PB1 genomes from human virus, NP, M and NS genomes from 'classical' swine virus, and PB2 and PA genomes from avian virus[38]. Then, the 'triple reassortant' swine descendant virus reassorted with the European avian-like porcine virus by acquiring the NA and M genomes to become A(H1N1) pdm2009 virus[39–41]. Thus, avian-origin PB2 and PA genes have co-evolved in the genetic backgrounds of swine IAVs. Although there are no pig- or human-specific mutations, higher numbers of I147T, A588T, and Q591R mutations in the Pig category than in the Human category, and different combinations of mutations, may reflect the unique mutational steps of avian-origin PB2 genes in swine[21]. As the MPM profiles of A(H1N1) pdm2009 and seasonal flu viruses are different, the early starting MPMs are likely to determine the accumulation patterns of subsequent mutations. However, the reasons why reassortant A(H1N1) pdm2009 viruses with stronger PB2 genes of seasonal flu viruses did not appear during their co-circulation may reflect importance of genetic compatibility of viral genomes[42,43].

Most PB2 mutations tested in this study increased polymerase activity, and the rank orders of the effects on polymerase activity and replication efficiency in the mouse lungs were identical in the cases of E627K, Q591K, E192K, K526R, A588I and Q591R. However, E158G, D253N, T271A and D701N showed different rank orders, and F404L, A588T and A674T did not increase polymerase activity. Therefore, we could not exclude the positive or negative epistasis of already-present mutations of 01310PB2 (661A and 683T) or other PR8 genes such as PA and PB1 genes[30,43]. As expected, E627K is most potent mutation for the mammalian pathogenicity of IAVs, and the high ranking of E627K, D701N and Q591K was coincident with their high frequency among quasi-species during the first mouse infection of a PR8-derived recombinant virus with the 01310 PB2-MVV gene[20,29,32,33]. The 591K mutation was much more potent than 591R in terms of mammalian pathogenicity. The high ranking but extremely low frequency of E158G mutation may raise concerns about a potent, potential risk factor for avian and swine PB2 genes if they evolve to the most compatible genomic constellation[10].
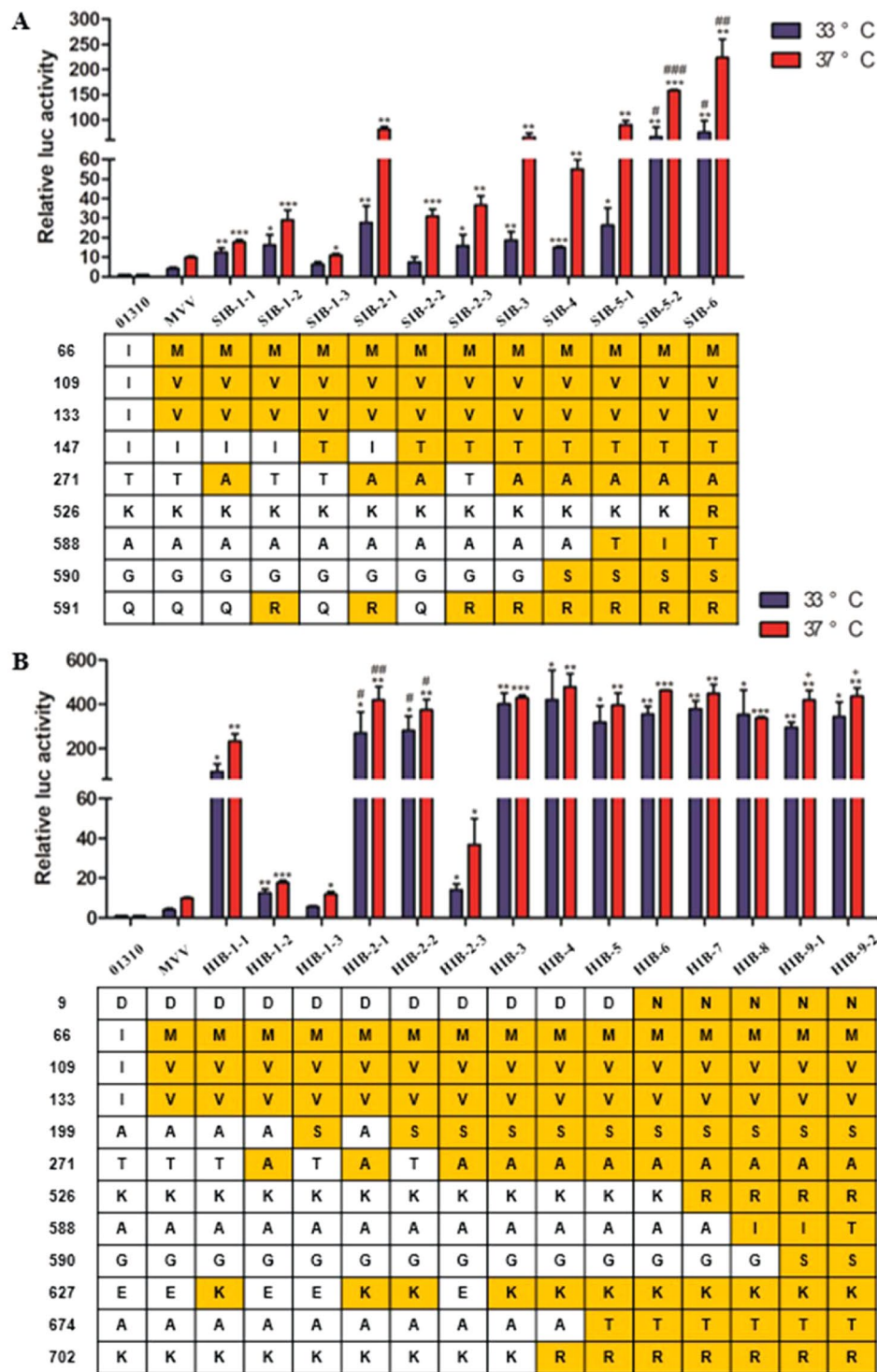
Figure 5. The effect of combined multiple MPMs on polymerase activity. Polymerase activity was measured using mini-genome assays in 293T cells at 33 and 37 °C. The data were normalized to the polymerase activity of the wild-type 01310 PB2 gene. Statistical significance was calculated using one-way ANOVA. Panel A, Polymerase activity of each combined MPM of the SIB-series (compared to PB2(01310)-MVV, $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$; compared to SIB-5-1, $^#P < 0.05$, $^{##}P < 0.01$, $^{###}P < 0.001$). The data are the average of three independent experiments ± s.d. Panel B, Polymerase activity of each combined MPMs of the HIB-series (compared to PB2(01310)-MVV, $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$; compared to HIB-1-1, $^#P < 0.05$, $^{##}P < 0.01$; compared to HIB-8, $^+ P < 0.05$). The data are the average of three independent experiments ± s.d.

A

B
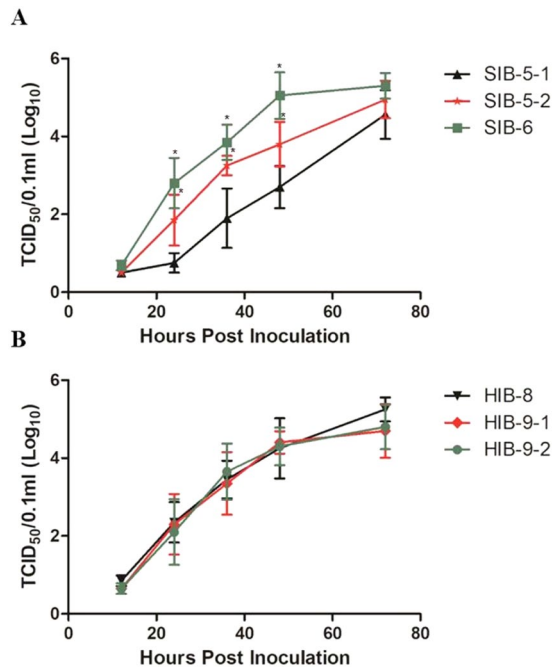
Figure 6. The effects of predominant swine (SIB-5-1, SIB-5-2 and SIB-6) and human (HIB-8, HIB-9-1 and HIB-9-2) PB2 genotypes on viral growth kinetics in A549 cells. The PR8-derived recombinant viruses containing major PB2 genotypes (panel A, SIB-series and panel B, HIB-series) were used to infect A549 cells (0.01 moi). The cells were incubated in a CO_2 incubator at 37 °C, and supernatants were collected at 12, 24, 36, 48, and 72 hpi to measure viral titers in MDCK cells ($TCID_{50}$). Statistical significance was analyzed by Student's t test (compared to SIB-5-1, $^{*}P < 0.05$). The data are the average of three independent experiments $\pm$ s.d.

PB2 interacts with different host factors to play multiple roles. The MPMs are located in different domains and may take part in different functions. The 627 domain containing closely located E627K, Q591K/R, G590S, and A/V588T/I mutations directly interacts with ANP32A to stimulate RNA synthesis, and E627K increases interaction with the mammalian isoform of ANP32A to result in high mammalian pathogenicity[6,9,44]. TUFM is another host restriction factor differentiating 627E from 627K and impedes the replication of avian IAVs with 627E via autophagy[7]. E627K and Q591R showed negative epistasis, and no synergistic effect was observed in combination[45]. This may be the reason for the extremely low frequency of E627K together with Q591R/K PB2 genes in all categories. However, other mutations such as G590S and A/V588T/I were acquired in addition to E627K or Q591R. The E627K mutation was not selected in avian but in mammalian hosts[29]. Therefore, acquisition of mutations in the 627 domain may be crucial for avian PB2 genes not only to replicate efficiently in mammalian hosts but also to acquire additional mutations for better fitness during competition between early mammal-adapted avian-origin PB2 genes.

The NLS domain interacts with importin-α, and the second highest effect of D701N on the replication efficiency in mouse lungs may reflect the importance of the nuclear localization of PB2[8]. However, unsuccessful acquisition of multiple mutations by PB2 genes with the first D701N mutation may imply that the first strategy to steeply increase nuclear localization of PB2 may not be sustainable. In addition, a significantly higher frequency of less active 702R (98.94%) relative to 701N (0.25%) in the Human category in PB2 bearing E627K may reflect a tendency to determine subsequent mutations by preceding mutations.

The exact function of the third highest effect of E158G in the lid domain is still unclear, but two additional mutations (E192K and A199S) in the lid domain may support the important role of the domain. The N-terminal 200 amino acid residues of PB2 are involved in the interaction with RNP to transcribe and replicate viral RNAs[46]. In addition, basic amino acids in the N-terminal half of PB2, such as 124R, 142R/143R, 268R, and 331K/332R, play important roles in both the transcription and replication of vRNAs[47]. The exact function of 147T is still unclear, but its close location to 142R/143R may be noteworthy[23]. The T271A mutation in the mid domain increased the replication efficiency of IAVs at the human upper respiratory tract temperature, 33-34°C, and individual (T271A) and combined (T271A with MVV) mutations significantly increased polymerase activity at 37°C or at both 33 and 37°C, respectively, relative to 01310PB2 in this study[24,30].

The role or function related to the highly ranked K526R mutation in the cap-627 linker domain is unclear, but the impinging position of 520-535 residues on the cap-binding site may support speculation regarding its role in viral mRNA transcription[48]. In addition, it is noteworthy that K526R showed a higher effect than the cap-binding site mutations on polymerase activity. Considering the report that K526R enhances the effects of E627K on viral replication via coordination with NEP (nuclear export protein), the role of K526R may not be as simple as has been postulated[14]. Although the polymerase activity of HIB-7 and HIB-8 was indistinguishable, the acquisition of the K526R mutation by rPB2(01310)-SIB-6 increased polymerase activity, replication efficiency in A549 cells
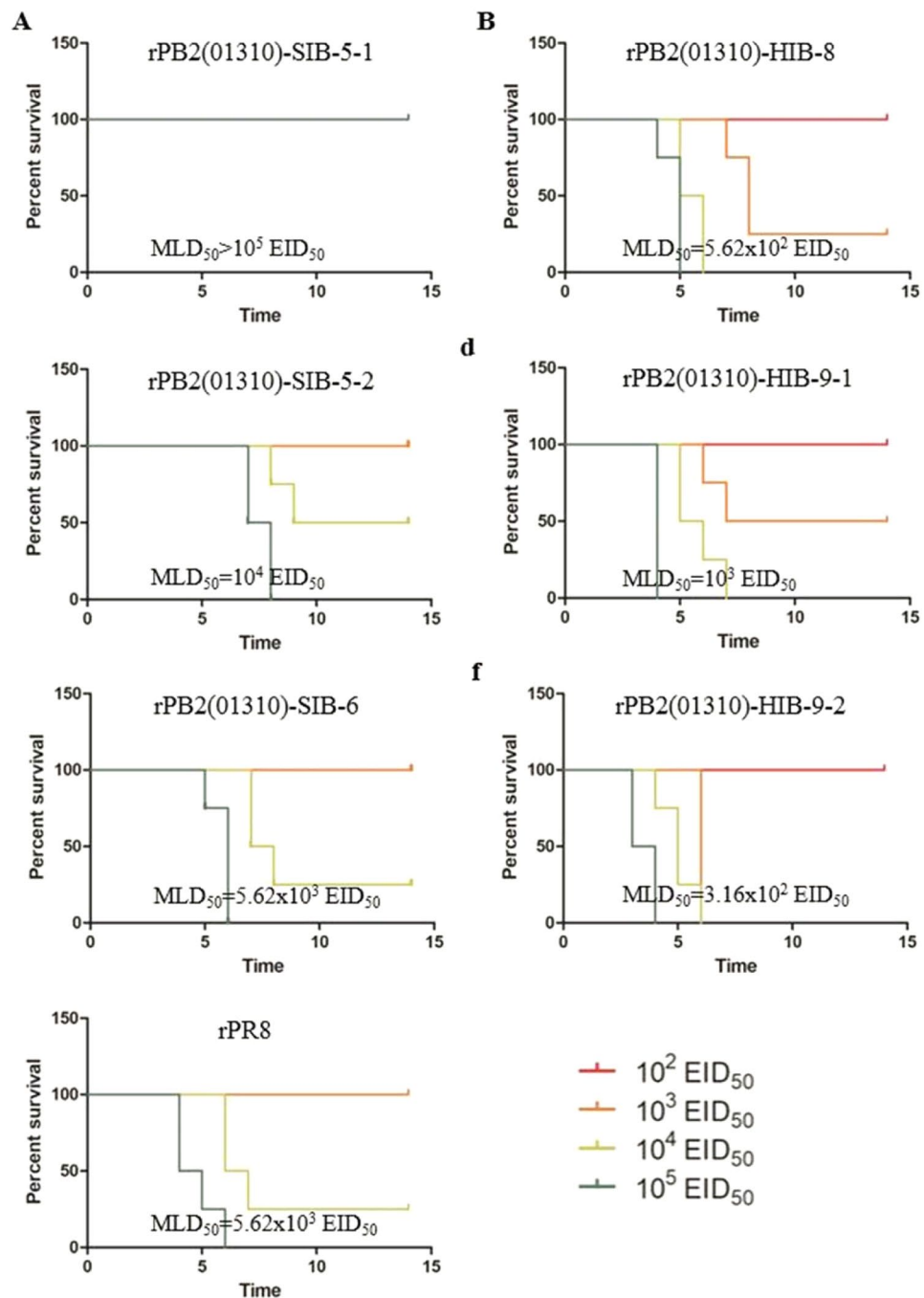
Figure 7. The effect of predominant swine and human PB2 genotypes on mouse pathogenicity. Six-week-old female BALB/c mice were intranasally inoculated with 10^2 to 10^5 EID_{50} of PR8-derived recombinant viruses containing major PB2 genotypes (panel A, SIB-series; panel B, HIB-series; panel C, rPR8), and MLD_{50} was measured. During the period up to 14 dpi, mice that lost more than 20% of their original weight were euthanized and recorded as dead.

and mouse pathogenicity in comparison with rPB2(01310)-SIB-5-1. The K526R mutation was suggested to be a mammalian adaptation marker of avian H7N9 viruses in collaboration with the E627K and D701N mutations and to play a role in the successful adaptation of H3N2 seasonal flu since approximately 1970[14].

The T588I mutation in the PB2 genes of A(H1N1) pdm2009 viruses increased polymerase activity at 33 and 37 °C and increased mouse pathogenicity by way of probable interaction with MAVS, resulting in inhibition of IFN-β expression[11]. The A588T mutation also increased polymerase activity and mouse pathogenicity, while 588I showed significantly higher polymerase activity than 588T in this study[23]. Although HIB-9-1 and HIB-9-2 did not show dramatic differences in polymerase activity and mouse pathogenicity, rPB2(01310)-SIB-5-2, which contained 588I, showed higher viral replication and pathogenicity in mammals than rPB2(01310)-SIB-5-1. The

high frequency of HIB-9-1 but recent dominance of HIB-9-2 may reflect the fine-tuning of mutation potency for viral fitness (I588T) as observed in mutational selection between 591K and 591R or 701N and 702R (Fig. 4). Interestingly, E627K was incompatible with the PB2 gene of A(H1N1) pdm2009, but most of its MPMs (147T, 590S, 591R, 271A, 526R and 588T/I) were successfully accumulated to E627K (Table S8)[45]. Therefore, the type of early MPMs may be important.

Previously, MVV mutation was speculated to be the minimum essential for avian PB2 genes to acquire additional mutations for mammalian adaptation[29]. The PB2 gene of A(H1N1) pdm1918 (A/Brevig Mission/1/1918) possessed only MVV, E627K, A199S, and K702R mutations, and they may be minimum essential PB2 mutations for catastrophic pandemics[29,49]. Although A(H1N1) pdm2009 viruses without strong PB2 mutations became pandemic, the low mouse pathogenicity of SIB-5-1 is coincident with the low pathogenicity of A(H1N1) pdm2009 viruses[50]. Early mutations in the 627 domain, e.g., G590S and E627K, may be crucial for viruses to become pandemic. A strong E627K mutation may be enough for viral fitness, but a weak G590S mutation may need to recruit more mutations, Q591R and A588T, in the 627 domain[51]. Although the D701N mutation is strong, it may be unsuccessful for recruiting additional mutations to become pandemic at present. However, PB2 genes possessing I147T-K339T mutations have evolved by acquiring E627K or A588T mutation, and H5N1 viruses carried those PB2 genes[23]. The high prevalence of I147T-K339T mutations in the avian category may mean they did not mitigate viral fitness in avian hosts, and they may need to be monitored for additional mutation acquisition and mammalian transmission (Fig. 4).

In this study, we collectively compared individual and combined MPMs in the same genetic backgrounds and determined their rank orders of effects on polymerase activity, replication efficiency in mammalian hosts and mouse pathogenicity. In addition, we provided hypothetical mutational steps of avian PB2 genes during mammalian adaptation. Thus, our data and the prototypic PB2 gene may be useful to predict present and future risks and to understand the evolution of avian and mammalian PB2 genes. Additionally, our present study may encourage more systematic studies to simulate co-evolution and epistatic interaction of PB2, PB1 and PA genes of IAVs.

Materials and Methods

Viruses, eggs, cells and plasmids.   A/Puerto Rico/8/34 (H1N1) (PR8) and PR8-derived viruses containing the mutated 01310 PB2 gene were generated using a Hoffmann's reverse genetic system[29,52,53]. They were passaged two times in 10-day-old SPF embryonated chicken eggs (ECEs) via the allantoic cavity route (Charles River Laboratories, Wilmington, MA, USA). 293T, MDCK, and A549 cells were purchased from the Korean Collection for Type Cultures (KCTC, Daejeon, Korea). 293T and MDCK cells were maintained in DMEM supplemented with 10% FBS (Life Technologies Co., Carlsbad, CA, USA), and A549 cells were maintained in DMEM/F12 supplemented with 10% FBS. To generate recombinant viruses, Hoffmann's reverse genetic system was used as described previously[52,53]. Recombinant viruses were generated and passaged two times in 10-day-old SPF ECEs and then stored at −80 °C until experimental use.

Data mining and sequence analysis.   PB2 genes containing a complete coding sequence were collected from the NCBI Influenza Virus Resource (http://www.fludb.org), and they were translated and compared with the BioEdit program (v7.2.5). The profiles of MPM genes were analyzed by median joining analysis with Network 5.0.0.3[54].

Cloning and site-directed mutagenesis.   The 01310PB2 segment previously cloned into Hoffmann's bidirectional transcription vector pHW2000 was used in this study[29]. The insert sequence was confirmed by sequencing with primers cmv-SF and bGH-SR as previously described (Macrogen Co, Seoul, Korea)[55]. Site-directed mutagenesis to introduce individual or combined MPMs into 01310 PB2 gene was performed using a Muta-direct Site Directed Mutagenesis Kit as per the manufacturer's protocol (iNtRON, Sungnam, Korea).

Mini-genome assay.   To evaluate the polymerase activity of mutated PB2 genes, we used the pHW-NP-Luc plasmid as previously described[29]. Briefly, 293T cells in 12-well plates were cotransfected with 0.1 μg each of pHW-NP-Luc and mutated 01310 PB2 and the PR8 PB1, PA and NP genes. Additionally, 0.1 μg of the *Renilla* luciferase plasmid pRL-TK (Promega, Madison, WI, USA) was also cotransfected, which served as an internal control to normalize variations in transfection efficiency and sample processing. Then, 24 hours after transfection, luminescence was measured using a Dual-Glo Luciferase Assay System (Promega, Madison, WI, USA) in accordance with the manufacturer's instructions on a TECAN Infinite 200 pro machine (Tecan Benelux bv, Giessen, Netherlands). All results shown are the average of triplicate experiments, and the standard deviation was calculated.

Generation of recombinant viruses by reverse genetics

PR8-derived recombinant virus was generated by transfecting Hoffmann's eight reverse genetics plasmids into 293T cells as described previously, with some modifications[52,53,55]. One day before the transfection, 293T cells were cultured in 6-well plates (5 × 10⁵ cells/well), and 300 ng of each plasmid was transfected together into the cells using Lipofectamine Plus reagent (Life Technologies) in 1 mL of Opti-MEM (Life Technologies). After overnight incubation, 1 mL of Opti-MEM and 0.5 μg/mL L-1-tosylamide-2-phenylethyl chloromethyl ketone (TPCK)-treated trypsin (Sigma-Aldrich, USA) were added. After another overnight incubation, 0.2 mL of the harvested culture medium was inoculated into 10-day-old SPF ECEs via the allantoic cavity route. The allantoic fluid was harvested from inoculated ECEs after incubation at 37 °C and checked for the presence and titers of recombinant viruses using plate and 96-well hemagglutination tests with 1% (v/v) chicken red blood

cells according to the WHO Manual on Animal Influenza Diagnosis and Surveillance. We verified nucleotide sequences of PB1 and PA genes of selected recombinant viruses as previous[55].

Titration of viruses. To estimate virus titers, each virus was serially diluted from 10^{-1} to 10^{-9} in 10-fold increments, and each dilution was injected into five 10-day-old SPF ECEs as well as inoculated onto MDCK cells. The 50% chicken embryo infectious dose (EID_{50}) and 50% tissue culture infectious dose ($TCID_{50}$) were calculated using the Spearman-Karber method[56].

Replication efficiency of recombinant viruses in MDCK cells. To evaluate the replication efficiency of each virus, MDCK cells (2×10^4/mL) were seeded in 96-well plates (100 μL/well). After 24 hours, confluent cells were washed twice with phosphate-buffered saline (PBS). Mutant viruses (10^7 EID_{50}/0.1 mL) were 10-fold serially diluted, and 200 μL of each dilution was inoculated into each well with fresh DMEM [supplemented with 1% bovine serum albumin (BSA) (fraction V) (Roche, Basel, Switzerland), 20 mM HEPES, antibiotic-antimycotic (Thermo Fisher Scientific, Waltham, MA, USA), and 1 μg/mL TPCK-treated trypsin (Sigma-Aldrich, St. Louis, MO, USA)]. During 5 days of incubation in a CO_2 incubator at 37 °C, supernatants were collected at 5 day-post-inoculation (dpi), and virus titers were determined as described above.

Growth kinetics of recombinant viruses in A549 cells. A549 cells were seeded in 12-well plates. After 24 hours, confluent cells were washed twice with PBS, and 0.01 multiplicity of infection (moi) of recombinant virus was inoculated into each well with 1 mL of fresh DMEM containing 0.25 μg/mL TPCK-treated trypsin. During 3 days of incubation in a CO_2 incubator at 37 °C, supernatant was harvested at 12, 24, 36, 48, and 72 hour-post-inoculation (hpi), and virus titers were determined as described above.

Animal experiments. Six-week-old female BALB/c mice (KOATEC Co. Pyeongtaek, Korea) were used for a mouse pathogenicity test (BIoPOA Co., Yongin, Korea). Five BALB/c mice were intranasally inoculated with 10^6 EID_{50}/50 μL of each virus after anaesthetized intraperitoneal injection of 15 mg/kg Zoletil 50 (Virbac, Carros, France). Negative control (mock) mice were injected with the same volume of sterilized PBS. To measure the 50% mouse lethal dose (MLD_{50}), four mice were anaesthetized via intraperitoneal injection of 15 mg/kg Zoletil 50 (Virbac, Carros, France) and then intranasally inoculated with 10^2 to 10^5 EID_{50}/50 μL of each virus. Mortality and weight loss were measured for 14 days. Mice that lost more than 20% of their original weight were euthanized and recorded as dead. MLD_{50} was calculated by the Spearman-Karber method. For the measurement of virus replication in infected mouse lung, six mice from each group were injected with PBS (mock) or 10^6 EID_{50}/50 μL of mutant virus. The lungs were collected at 3 and 6 dpi and then stored at -80 °C until use. The lungs were ground using a TissueLyzer 2 (Qiagen, Hilden, G) with 5 mM stainless steel beads and a volume of PBS equal to 10% of the lung weight in suspension. Then, 10 volumes of PBS were mixed with the ground tissues. After centrifugation at $2000 \times g$ for 10 min, the supernatants were used for viral titers, which were measured as EID_{50}.

Ethics statement. All mouse experiments were carried out at BioPOA Co. (Yongin, Korea) following a protocol that adhered to the National Institutes of Health's Public Health Service Policy on the Humane Care and Use of Laboratory Animals. The protocol was reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of BioPOA Co. (BP-2016-006-2).

Statistical analysis. The polymerase activity and virus titer were compared by one-way analysis-of-variance (IBM SPSS Statistics ver. 23; IBM, USA). The results were considered statistically significant if $P < 0.05$, $P < 0.01$ and $P < 0.001$.

Data availability
All data generated or analysed during this study are included in this published article (and its Supplementary Information files).

References
1. Webster, R. G. & Govorkova, E. A. Continuing challenges in influenza. *Ann. N. Y. Acad. Sci.* **1323**, 115–139, https://doi.org/10.1111/nyas.12462 (2014).
2. Guilligay, D. *et al.* The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat. Struct. Mol. Biol.* **15**, 500–506, https://doi.org/10.1038/nsmb.1421 (2008).
3. Hatta, M., Gao, P., Halfmann, P. & Kawaoka, Y. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. *Sci.* **293**, 1840–1842, https://doi.org/10.1126/science.1062882 (2001).
4. Subbarao, E. K., London, W. & Murphy, B. R. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *J. Virol.* **67**, 1761–1764 (1993).
5. Graef, K. M. *et al.* The PB2 subunit of the influenza virus RNA polymerase affects virulence by interacting with the mitochondrial antiviral signaling protein and inhibiting expression of beta interferon. *J. Virol.* **84**, 8433–8445, https://doi.org/10.1128/JVI.00879-10 (2010).
6. Baker, S. F., Ledwith, M. P. & Mehle, A. Differential Splicing of ANP32A in Birds Alters Its Ability to Stimulate RNA Synthesis by Restricted Influenza Polymerase. *Cell Rep.* **24**, 2581–2588 e2584, https://doi.org/10.1016/j.celrep.2018.08.012 (2018).
7. Kuo, S. M. *et al.* Inhibition of Avian Influenza A Virus Replication in Human Cells by Host Restriction Factor TUFM Is Correlated with Autophagy. *MBio* **8**, https://doi.org/10.1128/mBio.00481-17 (2017).
8. Gabriel, G., Herwig, A. & Klenk, H. D. Interaction of polymerase subunit PB2 and NP with importin alpha1 is a determinant of host range of influenza A virus. *PLoS Pathog* **4**, e11, https://doi.org/10.1371/journal.ppat.0040011 (2008).
9. Yamada, S. *et al.* Biological and structural characterization of a host-adapting amino acid in influenza virus. *PLoS Pathog.* **6**, e1001034, https://doi.org/10.1371/journal.ppat.1001034 (2010).

10. Zhou, B. *et al*. PB2 residue 158 is a pathogenic determinant of pandemic H1N1 and H5 influenza a viruses in mice. *J. Virol.* **85**, 357–365, https://doi.org/10.1128/JVI.01694-10 (2011).

11. Zhao, Z. *et al*. PB2-588I enhances 2009 H1N1 pandemic influenza virus virulence by increasing viral replication and exacerbating PB2 inhibition of beta interferon expression. *J. Virol.* **88**, 2260–2267, https://doi.org/10.1128/JVI.03024-13 (2014).

12. Liu, Q. *et al*. Virulence determinants in the PB2 gene of a mouse-adapted H9N2 virus. *J. Virol.* **89**, 877–882, https://doi.org/10.1128/JVI.01775-14 (2015).

13. Kim, J. H. *et al*. Role of host-specific amino acids in the pathogenicity of avian H5N1 influenza viruses in mice. *J. Gen. Virol.* **91**, 1284–1289, https://doi.org/10.1099/vir.0.018143-0 (2010).

14. Song, W. *et al*. The K526R substitution in viral protein PB2 enhances the effects of E627K on influenza virus replication. *Nat. Commun.* **5**, 5509, https://doi.org/10.1038/ncomms6509 (2014).

15. Czudai-Matwich, V., Otte, A., Matrosovich, M., Gabriel, G. & Klenk, H. D. PB2 mutations D701N and S714R promote adaptation of an influenza H5N1 virus to a mammalian host. *J. Virol.* **88**, 8735–8742, https://doi.org/10.1128/JVI.00422-14 (2014).

16. Mok, C. K. *et al*. Amino acid residues 253 and 591 of the PB2 protein of avian influenza virus A H9N2 contribute to mammalian pathogenesis. *J. Virol.* **85**, 9641–9645, https://doi.org/10.1128/JVI.00702-11 (2011).

17. Shinya, K. *et al*. PB2 amino acid at position 627 affects replicative efficiency, but not cell tropism, of Hong Kong H5N1 influenza A viruses in mice. *Virology* **320**, 258–266, https://doi.org/10.1016/j.virol.2003.11.030 (2004).

18. Finkelstein, D. B. *et al*. Persistent host markers in pandemic and H5N1 influenza viruses. *J. Virol.* **81**, 10292–10299, https://doi.org/10.1128/JVI.00921-07 (2007).

19. Li, Z. *et al*. Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. *J. Virol.* **79**, 12058–12064, https://doi.org/10.1128/JVI.79.18.12058-12064.2005 (2005).

20. Steel, J., Lowen, A. C., Mubareka, S. & Palese, P. Transmission of influenza virus in a mammalian host is increased by PB2 amino acids 627K or 627E/701N. *PLoS Pathog.* **5**, e1000252, https://doi.org/10.1371/journal.ppat.1000252 (2009).

21. Mehle, A. & Doudna, J. A. Adaptive strategies of the influenza virus polymerase for replication in humans. *Proc. Natl Acad. Sci. USA* **106**, 21312–21316, https://doi.org/10.1073/pnas.0911915106 (2009).

22. Taft, A. S. *et al*. Identification of mammalian-adapting mutations in the polymerase complex of an avian H5N1 influenza virus. *Nature. Commun.* **6**, 7491, https://doi.org/10.1038/ncomms8491 (2015).

23. Fan, S. *et al*. Novel residues in avian influenza virus PB2 protein affect virulence in mammalian hosts. *Nat. Commun.* **5**, 5021, https://doi.org/10.1038/ncomms6021 (2014).

24. Bussey, K. A., Bousse, T. L., Desmet, E. A., Kim, B. & Takimoto, T. PB2 residue 271 plays a key role in enhanced polymerase activity of influenza A viruses in mammalian host cells. *J. Virol.* **84**, 4395–4406, https://doi.org/10.1128/jvi.02642-09 (2010).

25. Liu, Q. *et al*. Combination of PB2 271A and SR polymorphism at positions 590/591 is critical for viral replication and virulence of swine influenza virus in cultured cells and *in vivo*. *J. Virol.* **86**, 1233–1237, https://doi.org/10.1128/JVI.05699-11 (2012).

26. Danzy, S. *et al*. Mutations to PB2 and NP proteins of an avian influenza virus combine to confer efficient growth in primary human respiratory cells. *J. Virol.* **88**, 13436–13446, https://doi.org/10.1128/JVI.01093-14 (2014).

27. Resa-Infante, P. *et al*. The host-dependent interaction of alpha-importins with influenza PB2 polymerase subunit is required for virus RNA replication. *PLoS One* **3**, e3904, https://doi.org/10.1371/journal.pone.0003904 (2008).

28. Lyons, D. M. & Lauring, A. S. Mutation and Epistasis in Influenza Virus Evolution. *Viruses* **10**, https://doi.org/10.3390/v10080407 (2018).

29. Lee, C. Y. *et al*. Prerequisites for the acquisition of mammalian pathogenicity by influenza A virus with a prototypic avian PB2 gene. *Sci. Rep.* **7**, 10205, https://doi.org/10.1038/s41598-017-09560-z (2017).

30. Hayashi, T., Wills, S., Bussey, K. A. & Takimoto, T. Identification of Influenza A Virus PB2 Residues Involved in Enhanced Polymerase Activity and Virus Growth in Mammalian Cells at Low Temperatures. *J. Virol.* **89**, 8042–8049, https://doi.org/10.1128/JVI.00901-15 (2015).

31. Zhang, H. *et al*. The PB2 E627K mutation contributes to the high polymerase activity and enhanced replication of H7N9 influenza virus. *J. Gen. Virol.* **95**, 779–786, https://doi.org/10.1099/vir.0.061721-0 (2014).

32. Min, J. Y. *et al*. Mammalian adaptation in the PB2 gene of avian H5N1 influenza virus. *J. Virol.* **87**, 10884–10888, https://doi.org/10.1128/JVI.01016-13 (2013).

33. Subbarao, E. K., London, W. & Murphy, B. R. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *J. Virol*, 1764–1764 (1993).

34. Zhu, W. *et al*. Dual E627K and D701N mutations in the PB2 protein of A(H7N9) influenza virus increased its virulence in mammalian models. *Sci. Rep.* **5**, 14170, https://doi.org/10.1038/srep14170 (2015).

35. Mok, C. K. *et al*. Amino acid substitutions in polymerase basic protein 2 gene contribute to the pathogenicity of the novel A/H7N9 influenza virus in mammalian hosts. *J. Virol.* **88**, 3568–3576, https://doi.org/10.1128/JVI.02740-13 (2014).

36. Sanjuan, R., Nebot, M. R., Chirico, N., Mansky, L. M. & Belshaw, R. Viral mutation rates. *J. Virol.* **84**, 9733–9748, https://doi.org/10.1128/JVI.00694-10 (2010).

37. Shope, R. E. Swine Influenza: I. Experimental Transmission and Pathology. *J. Exp. Med.* **54**, 349–359 (1931).

38. Webby, R. J., Rossow, K., Erickson, G., Sims, Y. & Webster, R. Multiple lineages of antigenically and genetically diverse influenza A virus co-circulate in the United States swine population. *Virus Res.* **103**, 67–73, https://doi.org/10.1016/j.virusres.2004.02.015 (2004).

39. Bourret, V. Avian influenza viruses in pigs: An overview. *Vet. J.* **239**, 7–14, https://doi.org/10.1016/j.tvjl.2018.07.005 (2018).

40. Garten, R. J. *et al*. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Sci.* **325**, 197–201, https://doi.org/10.1126/science.1176225 (2009).

41. Smith, G. J. *et al*. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nat.* **459**, 1122–1125, https://doi.org/10.1038/nature08182 (2009).

42. Phipps, K. L. *et al*. Seasonal H3N2 and 2009 Pandemic H1N1 Influenza A Viruses Reassort Efficiently but Produce Attenuated Progeny. *J Virol* **91**, https://doi.org/10.1128/JVI.00830-17 (2017).

43. Chen, K. Y., Santos Afonso, E. D., Enouf, V., Isel, C. & Naffakh, N. Influenza virus polymerase subunits co-evolve to ensure proper levels of dimerization of the heterotrimer. *PLoS Pathog.* **15**, e1008034, https://doi.org/10.1371/journal.ppat.1008034 (2019).

44. Tarendeau, F. *et al*. Host determinant residue lysine 627 lies on the surface of a discrete, folded domain of influenza virus polymerase PB2 subunit. *PLoS Pathog.* **4**, e1000136, https://doi.org/10.1371/journal.ppat.1000136 (2008).

45. Jagger, B. W. *et al*. The PB2-E627K mutation attenuates viruses containing the 2009 H1N1 influenza pandemic polymerase. *MBio* **1**, https://doi.org/10.1128/mBio.00067-10 (2010).

46. Kashiwagi, T. *et al*. The N-terminal fragment of a PB2 subunit from the influenza A virus (A/Hong Kong/156/1997 H5N1) effectively inhibits RNP activity and viral replication. *PLoS One* **9**, e114502, https://doi.org/10.1371/journal.pone.0114502 (2014).

47. Hara, K., Kashiwagi, T., Hamada, N. & Watanabe, H. Basic amino acids in the N-terminal half of the PB2 subunit of influenza virus RNA polymerase are involved in both transcription and replication. *J. Gen. Virol.* **98**, 900–905, https://doi.org/10.1099/jgv.0.000750 (2017).

48. Ma, X. *et al*. Structural basis for therapeutic inhibition of influenza A polymerase PB2 subunit. *Sci. Rep.* **7**, 9385, https://doi.org/10.1038/s41598-017-09538-x (2017).

49. Taubenberger, J. K. *et al*. Characterization of the 1918 influenza virus polymerase genes. *Nat.* **437**, 889–893, https://doi.org/10.1038/nature04230 (2005).

50. Novel Swine-Origin Influenza, A. V. I. T. *et al.* Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N. Engl. J. Med.* **360**, 2605–2615, https://doi.org/10.1056/NEJMoa0903810 (2009).
51. Wang, C., Lee, H. H., Yang, Z. F., Mok, C. K. & Zhang, Z. PB2-Q591K Mutation Determines the Pathogenicity of Avian H9N2 Influenza Viruses for Mammalian Species. *PLoS One* **11**, e0162163, https://doi.org/10.1371/journal.pone.0162163 (2016).
52. Hoffmann, E., Neumann, G., Kawaoka, Y., Hobom, G. & Webster, R. G. A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc. Natl Acad. Sci. USA* **97**, 6108–6113, https://doi.org/10.1073/pnas.100133697 (2000).
53. Hoffmann, E., Krauss, S., Perez, D., Webby, R. & Webster, R. G. Eight-plasmid system for rapid generation of influenza virus vaccines. *Vaccine* **20**, 3165–3170 (2002).
54. Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254–267, https://doi.org/10.1093/molbev/msj030 (2006).
55. Kim, I. H., Choi, J. G., Lee, Y. J., Kwon, H. J. & Kim, J. H. Effects of different polymerases of avian influenza viruses on the growth and pathogenicity of A/Puerto Rico/8/1934 (H1N1)-derived reassorted viruses. *Vet. Microbiol.* **168**, 41–49, https://doi.org/10.1016/j.vetmic.2013.10.011 (2014).
56. Hamilton, M. A., Russo, R. C. & Thurston, R. V. Trimmed Spearman-Karber method for estimating median lethal concentrations in toxicity bioassays. *Environ. Sci. Technol.* **11**, 714–719 (1977).

Acknowledgements

Author contributions

H.-J.K. and C.-Y.L. designed the study, analysed the data and wrote the manuscript. C.-Y.L. and S.-H.A. performed experiments. J.-H.K., Y.-J.L. and J.-G.C. contributed data interpretation. Y.-J.L. and J.-G.C. provided the 01310 viruses and genetic information.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-62036-5.

Correspondence and requests for materials should be addressed to H.-J.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.