

Title

Molecular dissection of the structural and nonstructural proteins of spanish-1918

Influenza, pandemic-2009, and bird flu viruses

Authors

Gusti Ngurah Mahardika,^{1*} Nyoman Suartha,² Gusti Ayu Yuniati Kencana,¹ Ida

Bagus Kade Suardana,¹ Nyoman Sri Budayanti³

Author affiliations

¹Virology Laboratory, Faculty of Veterinary Medicine Udayana University, Denpasar, Bali, Indonesia

²Animal Hospital, Faculty of Veterinary Medicine Udayana University, Denpasar, Bali, Indonesia

³Microbiology Department, Faculty of Medicine Udayana University, Denpasar, Bali, Indonesia

***Corresponding Author**

Gusti Ngurah Mahardika, Virology Laboratory, Faculty of Veterinary Medicine

Udayana University, Jl. PB Sudirman, Denpasar, Bali, Indonesia. Phone/FAX: +62

361223791; Email: gnmahardika@unud.ac.id

Abstract

The potential emergence of deadly pandemic influenza viruses is unpredictable and most have emerged with no forewarning. The distinct epidemiological and pathological patterns of the Spanish (H1N1), pandemic-2009 (H1N1), and avian influenza (H5N1), known as bird flu, viruses may allow us to develop a ‘template’ for possible emergence of devastating pandemic strains. Here, we provide a detailed molecular dissection of the structural and nonstructural proteins of this triad of viruses. GenBank data for three representative strains were analyzed to determine the polymorphic amino acids, genetic distances, and isoelectric points, hydrophobicity plot, and protein modeling of various proteins. We propose that the most devastating pandemic strains may have full-length PB1-F2 protein with unique residues, highly cleavable HA, and a basic NS1. Any newly emerging strain should be compared with these three strains, so that resources can be directed appropriately.

Keywords:

Influenza; Spanish-1918; pandemic-2009; bird flu; avian influenza H5N1; superimposed hydrophobicity plot; protein modeling;

Background

We must be constantly alert to the potential emergence of deadly pandemic influenza viruses because most have emerged with no forewarning. The emergence of new strains will continue to challenge public health and pose problems for scientific communities¹. The distinct epidemiological patterns of the Spanish (H1N1), pandemic-2009 (H1N1), and bird flu (H5N1) viruses should allow us to develop a ‘template’ for possibly devastating pandemic strains. Spanish flu and pandemic-2009 flu established human-to-human infections^{2,3}, whereas bird flu, the highly pathogenic avian influenza virus (HPAIV) subtype H5N1 (HPAIV-H5N1), has so far shown only limited or unsustainable human-to-human transmission⁴⁻⁶. These three strains have spread worldwide, although bird flu has yet to arrive on the American continent⁷, while this is still theoretically possible⁸.

Spanish flu, which caused an estimated 20–50 million deaths around 1918⁹, is recognized as the most devastating epidemic in modern history¹⁰. However, the case fatality rate (CFR) of Spanish flu is not the highest recorded, although its global CFR is believed to have been around 2.5%³. The CFR of pandemic-2009 is estimated to have been approximately 0.05% of confirmed cases, so its virulence is deemed mild¹¹. The CFR increases greatly if the denominator is influenza-related critical illness. The mortality rate associated with critical illness in pandemic-2009 in various countries was reported to range from 15% in Australia to 61% in Southeast Asia¹². Among the viral triad, the CFR of bird flu is highest. As of December 19, 2016, the cumulative number of confirmed human cases of avian influenza A (H5N1) reported to the World Health Organization in 2003–2006 was 856, 452 (52.8%) of which were

fatal (www.who.int). However, this comparison of mortality rates is invalid because the availability of and access to medical interventions were significantly different in 1918 from now. The fatalities associated with pandemic-2009 and bird flu may have been much higher without modern medical modalities.

The influenza viruses are a unique family of viruses, the *Orthomyxoviridae*, and members have an eight-segment negative-strand genome¹³. Five segments encode single proteins: (from the largest to smallest) polymerase basic 2 (PB2), polymerase acidic (PA), hemagglutinin (HA) nucleoprotein (NP), and neuraminidase (NA)¹³. The second largest fragment encodes polymerase basic 1 (PB1)¹³ and an accessory peptide, PB1-F2^{14,15}. The two smallest segments are each spliced to produce the mRNAs for two proteins, M1 and M2 (the seventh segment) and NS1 and NS2 (the eighth segment)¹³.

No direct molecular comparison of all the genes of this influenza virus triad has been made. Molecular modeling of NA has been reported, which established the susceptibility patterns of the viruses to anti-influenza drugs¹⁶. A phylogenetic analysis and amino acid homology analysis of the polymerase complex genes of Spanish flu, seasonal flu, and bird flu have been described, and concluded that the genes of Spanish Flu closely resemble those of bird flu¹⁷. Although the analysis of viral genomes alone is unlikely to clarify some critical issues³, such as the viral capacity for human-to-human transmission and the severity of clinical infection, a head-to-head molecular comparison of different lineages should pinpoint the putative gene(s) or protein(s) that can be used as signals for the potential emergence of a catastrophic pandemic strain. Here, we provide a detailed molecular analysis of the

structural and nonstructural proteins of this triad of viruses that are associated with their distinct transmissibility and severity.

Results

The genetic distances, numbers of amino acid differences, and isoelectric points of all the structural and nonstructural proteins of the SF, PDM, and BF viruses are presented in Table 2. The lengths of all the corresponding proteins of the triad were equal, except those of PB1-F2, NS1, and HA. In SF and BF, PB1-F2 contains 90 residues, whereas in PDM, it contains only 11 residues. NS1 in PDM has lost 11 amino acids at the carboxyl terminus. The length of HA in SF and PDM is 566 amino acids, whereas in BF, it is 568 amino acids. The genetic distance between BF and SF is smaller than that between PDM and SF based on PB1 and PA, whereas the contrary true when the distances are based on PB2, HA, NS1, or NS2. The genetic distances are almost equal between BF and SF and between PDM and SF when based on other protein-coding fragments. Fewer amino acid differences are in all protein of BF to SF than PDM to SF, except in HA, NS1, and NS2. The isoelectric points of all proteins are almost equal in the three viruses, except that HA of PDM and NS1 of SF are more basic than the corresponding proteins in the other viruses.

The superimposed hydrophobicity plots of all the proteins are given in Supplementary Material 1. The plots of PB1-F2, HA, NS1, and NS2 of SF, PDM, and BF are presented in Figure 1. The plots for PB2, PB1, PA, NP, NA, MA1, and MA2 were almost perfectly superimposed. A slight aberration was observed in HA, which is

neutral in SF and PDM, but hydrophilic in BF at residues 300–400 (Figure 1). PB1-F2 is also more hydrophilic in BF than in SF at positions 1–20 and 60–70, but more hydrophobic at positions 30–50. The NS1 and NS2 plots were not superimposed exactly, and the mismatch was most prominent in the BF plot (Figure 1).

The result of protein modeling of PB1-F2, NS1, and NS2 are presented in Figures 2–4. PB1-F2 and NS2 are dominated by an α -helix, with no β -sheet. NS1 consists of an α -helix membrane domain and a globular head consisting of an α -helix and a β -sheet. All those viral proteins have homologous structures, with some minor differences. PB1-F2 consists of two α -helices. The amino-terminal helix is continuous in BF, whereas it is interrupted in SF. On the contrary, the carboxyl-terminal helix of SF is continuous, whereas that of BF is interrupted by random coils. NS1 consists of a globular head formed by the amino terminus of the protein, and a trans-membrane domain at the carboxyl terminus. NS2 consists of four α -helix structures. The first helix (from the amino terminus) is intact in BF, but interrupted in SF and PDM. The second (yellow in Figure 4) is intact in SF and BF, but interrupted in PDM. The third (green to light blue in Figure 4) is intact in all three viruses. The last helix at the carboxyl terminus (blue in Figure 4) is intact in SF, but interrupted in PDM and BF.

Discussion

It is generally believed that the pathogenicity and transmissibility of influenza viruses are polygenic or multifactorial^{18,19}. Many gene segments of these viruses contribute to their capacity to cause severe outcomes in their host and to be readily transmitted between hosts²⁰. Whereas one viral protein might define a pathogenic characteristic,

virulent outcome may only be possible with the contribution of other proteins. The magnitude of an infection depends on those proteins acting in concert with host and environmental factors²¹⁻²⁴. In other word, the genetic make-up of the virus must match the permissiveness of the host and an appropriate environment to generate fatal or contagious outcomes.

The full genomic sequence of the Spanish-1918 influenza virus, which was derived from a naturally preserved human body believed to have died in that severe pandemic^{17,25}, and the massive number of influenza virus sequences determined in recent decades provide data from which a ‘template’ for the emergence of modern disastrous pandemic strains can be generated.

The triad strains SF, PDM, and BF are representative of different influenza viruses with distinct hallmarks. All have been disseminated globally, except BF, which so far affects only Asia, Europe, and Africa⁷. The SF and PDM viruses are readily transmitted between human, whereas BF has no capacity for sustained human-to-human transmission⁴⁻⁶. BF has the highest CFR, at >50% (www.who.int). Without modern medical resources, as occurred in 1918, its CFR might be much higher. Spanish-1918 influenza, responsible for the ‘mother of pandemics’, had a CFR of 2%–3%³. Compared with those viruses, the pandemic-2009 virus is ‘avirulent’ or mild¹¹, with a CFR of only 0.05%¹¹.

We analyzed the genetic distances between the triad viruses, the differences in the numbers of amino acids and the isoelectric points of all their structural and nonstructural proteins, as well as their hydrophobicity plots. We found that BF closely

resembles SF, except in proteins HA and NS2. In those proteins, PDM is closer to SF than is BF. We identified PB1-F2, HA, and NS1 are the factors putatively responsible for the distinct human-to-human transmissibility of SF, PDM, and BF, and as hallmarks of disease severity.

PB1-F2 seems to be responsible for pathogenicity. PB1-F2 contains 90 residues in SF and BF, but only 11 residues in PDM. Therefore, the mild disease associated with PDM might be related to the loss of PB1-F2. Different hydrophobicity patterns might also be responsible for the differences between SF and BF. PB1-F2 is more hydrophilic at positions 1–20 and 60–70 in BF than in SF, but more hydrophobic at positions 30–50.

The obvious differences in the open reading frame (ORF) length of this peptide in SF, PDM, and BF suggest that PB1-F2 is responsible for the pathological outcomes of influenza viral infections. The truncation of PB1-F2 in PDM might have caused the low CFR of PDM in 2009. PB1-F2 is a small accessory protein encoded by an alternative ORF in the second largest segment of most influenza A virus genomes²⁶ and is thought to contribute to viral pathogenicity and the severity of pandemic influenza¹⁵.

The role of PB1-F2 in the pathogenesis of the influenza viruses is contentious. The function of this peptide is both strain- and host-specific²⁷, and it is even expressed strain-specifically²⁸. In various studies, this protein caused immune disruption through the recruitment of neutrophils and the inhibition of natural killer cells²⁹, immune cell apoptosis²⁶, and the inhibition of interferon³⁰. Its species-specific

activity was demonstrated in the attenuation of AIV-H5N1³¹ and in an experiment with swine isolates²⁸. The posttranslational phosphorylation of Ser³⁵³² and Ser⁶⁶³³ of PB1-F2 may contribute to the strain-specific functions this protein. Both SF and BF have serine at residue 35, whereas Ser⁶⁶ in SF is substituted with asparagine in BF.

The hydrophilic domain around the cleavage site of HA may be responsible for the highly pathogenic nature of BF. HA contains 566 amino acids in SF and PDM, but 568 amino acids in BF. The superimposition of the hydrophobicity plots of HA showed slight aberrations, with more hydrophilic amino acids in the region defined by residues 300–400 in BF than in the other viruses. This domain includes the cleavage site of HA.

HA is an important surface protein for receptor recognition and the penetration of the virus to the host cell cytoplasm¹³. Cleavage of this virus is a prerequisite for viral infections and is therefore a crucial determinant in viral pathogenicity and tissue tropism³⁴. The most highly pathogenic strains have polybasic amino acids at the cleavage site or no carbohydrate residues in the vicinity of the site¹⁸. The length of the site also enhances its cleavability¹⁸. The long track of basic amino acids at the cleavage site of HA in HPAIVs is thought to facilitate the expansion of tissue tropism. The cleavage of the HA precursor to HA1 and HA2 is mediated by a ubiquitous endopeptidase, furin, located in the trans-Golgi network¹⁹. Infectious virions are released from infected cells with no requirement for extracellular activation. Moreover, the cleavability and pathogenicity of HA seem to be strain- and host-specific. Zhang et al. showed experimentally that a single substitution in the cleavage site of HA modulates the virulence of the H5N1 virus³⁵, and another report³⁶

demonstrated that the contribution of the H5 multibasic site to the virulence of the HPAIV-H5N1 virus varies among mammalian hosts, and is most significant in mice and ferrets and less remarkable in nonhuman primates. The postulate of Horimoto and Kawaoka³⁷ that the cleavability of HA correlates with the degree of virulence, when all other genetic characteristics are considered equal, is valid. However, in the evolution of a low-pathogenic avian influenza virus to become an HPAIV, a change in the cleavage site alone is not enough. The low-pathogenic strains may already have cryptic virulence potential³⁸.

NS1 and NS2 may be responsible for the human-to-human transmission of the influenza virus. The PDM virus resembles SF more closely than BF. The hydrophobicity plots of NS1 and NS2 were incompletely superimposed, and the aberration was predominantly located in the BF plot (Figure 1). NS1 and NS2 are encoded by the shortest segment of the influenza genome. NS1 is a nonstructural protein expressed in high abundance in infected cells, whereas NS2 seems to be a structural protein that is a minor component of the virion¹³. NS1 is collinearly translated from the transcript of the segment, whereas NS2 is encoded by a spliced transcript¹³. NS1 is particularly linked to the ‘cytokine storm’ phenomenon that follows influenza infection³⁹, contributing to its unusual immunopathogenesis⁴⁰. The hyperinduction of proinflammatory cytokines is a hallmark of severe influenza infection⁴¹. The presence of host cofactors determines the effect of this protein. NS1 only induces a cytokine storm in the presence of the myeloperoxidase system³⁹, but the mechanism of NS2 still requires clarification. NS1 of SF is remarkably basic, whereas NS1 of the other viruses is acidic (Table 2). This unique property of SF NS1 might have led to the devastating humanitarian impact of the SF pandemic shortly before the 1920s.

Amino acid differences disturb of secondary structures of PB1, NS1, and NS2, which might alter the human-to-human transmissibility and human pathogenicity of this triad of viruses. The modeling of the PB1-F2, NS1, and NS2 proteins showed that the proteins of all three viruses have homologous structures, with some minor differences.

The findings of this study are merely hypothetical. Reverse genetic experiments with combined gene segments are the only way to validate our hypotheses. Such experiments will be controversial and should be strictly regulated. A wide survey of the influenza virus genomes available in databases will offer indirect evidence to support our findings. The analysis conducted in this study was very simple and could be undertaken in many countries, so a capacity to immediately predict the potential impact of an emergent strain is possible in these countries.

We conclude that the putative pathogenicity of an influenza virus lies in PB1-F2 and the cleavability of HA, whereas NS1 and NS2 (especially NS1) are responsible for the human permissiveness of the virus. The most devastating pandemic strains may have full-length PB1-F2 proteins with unique residues, highly cleavable HA, and a basic NS1. The generation of such strain with reverse genetics will provide proof of this model. However, this kind of experiment must be strictly regulated or may even be impossible. Any newly emerging strain should be compared with this triad of influenza viruses to rapidly estimate its pathogenicity and human-to-human transmissibility.

Materials and Methods

The nucleotide and deduced amino acid sequences of all the structural and nonstructural proteins (PB2, PB1, PB1-F2, HA, NP, NA, M1, M2, NS1, and NS2) of Spanish Flu, pandemic-2009, and bird flu were analyzed in this study. Strain A/Brevig Mission/1/1918(H1N1) provided all the genomic segments of Spanish flu, except the HA gene, and strain A/South Carolina/1/18 (H1N1) provided the HA gene. The A/California/04-061-MA/2009(H1N1) strain represented the pandemic-2009 virus, and strain A/goose/Guangdong/1/1996(H5N1) represented the bird flu virus. The viruses are abbreviated hereinafter to SF, PDM, and BF for Spanish flu, pandemic-2009, and bird flu, respectively. The accession numbers for the GenBank data are presented in Table 1. The sequences were aligned using ClustalW, and the polymorphic amino acids in each protein and the numbers of amino acid differences were determined with Mega 6.0⁴². Genetic distances were calculated with the Kimura 2-parameter model⁴³. The isoelectric points were calculated with Peptide Calculator (<http://www.bachem.com>). A hydrophobicity plots of each peptide was constructed with the Kyte and Doolittle method⁴⁴, using the online software ProtScale⁴⁵ (<http://web.expasy.org/protscale/>), and the results were superimposed using Adobe Creative Cloud 2015 (Adobe System Incorporated). Protein modeling was performed with the online resource PYRE2 (<http://www.sbg.bio.ic.ac.uk>)⁴⁶, and the protein models were visualized with RasWin 2.7.5.2 (www.rasmol.org).

Acknowledgments

We thank Mr. Wisnu Wira Mahardika, a student at Bina Nusantara University in Tangerang–Banten, Indonesia, for assistance in generating the computer graphics. All authors are member of research team funded by Professor Publication and Promotion Project Udayana University of Bali, DIPa-PNBP 2018, Grant No. **383-1/UN14.4.A/LT/2018** with GNM as Principal Investigator. GAYK, IBKS and GNM are member of research team

funded by The National Key Research and Development Program of China, Grant No.

2016YFE0203200; in which GNM is Indonesian Principal Investigator. The funders have no role in drafting and finalizing the document and in making decision to publish.

Additional information

Compliance with Ethics Guidelines

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest

None

Author contributions

GNM and NSB conceived of or designed study. NSB, NS, GAYK and IBKS performed data collection. All author analyzed data. GNM and NSB wrote the paper.

Reference

- 1 Neumann, G., Noda, T. & Kawaoka, Y. Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature* **459**, 931-939, doi:10.1038/nature08157 (2009).
- 2 Mena, I. *et al.* Origins of the 2009 H1N1 influenza pandemic in swine in Mexico. *Elife* **5**, doi:10.7554/eLife.16777 (2016).
- 3 Taubenberger, J. K. & Morens, D. M. 1918 Influenza: the mother of all pandemics. *Emerging infectious diseases* **12**, 15-22, doi:10.3201/eid1201.050979 (2006).
- 4 Peiris, J. S., de Jong, M. D. & Guan, Y. Avian influenza virus (H5N1): a threat to human health. *Clinical microbiology reviews* **20**, 243-267, doi:10.1128/CMR.00037-06 (2007).

- 5 Kandun, I. N. *et al.* Three Indonesian clusters of H5N1 virus infection in 2005. *The New England journal of medicine* **355**, 2186-2194, doi:10.1056/NEJMoa060930 (2006).
- 6 Wang, H. *et al.* Probable limited person-to-person transmission of highly pathogenic avian influenza A (H5N1) virus in China. *Lancet* **371**, 1427-1434, doi:10.1016/S0140-6736(08)60493-6 (2008).
- 7 Li, X. *et al.* Global and local persistence of influenza A(H5N1) virus. *Emerging infectious diseases* **20**, 1287-1295, doi:10.3201/eid2008.130910 (2014).
- 8 Kilpatrick, A. M. *et al.* Predicting the global spread of H5N1 avian influenza. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 19368-19373, doi:10.1073/pnas.0609227103 (2006).
- 9 Tumpey, T. M. *et al.* Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **310**, 77-80, doi:10.1126/science.1119392 (2005).
- 10 Trilla, A., Trilla, G. & Daer, C. The 1918 "Spanish flu" in Spain. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **47**, 668-673, doi:10.1086/590567 (2008).
- 11 Nishiura, H. The virulence of pandemic influenza A (H1N1) 2009: an epidemiological perspective on the case-fatality ratio. *Expert Rev Respir Med* **4**, 329-338, doi:10.1586/ers.10.24 (2010).
- 12 Duggal, A., Pinto, R., Rubenfeld, G. & Fowler, R. A. Global Variability in Reported Mortality for Critical Illness during the 2009-10 Influenza A(H1N1) Pandemic: A Systematic Review and Meta-Regression to Guide Reporting of Outcomes during Disease Outbreaks. *PloS one* **11**, e0155044, doi:10.1371/journal.pone.0155044 (2016).
- 13 Lamb, R. A. & Krug, R. M. in *Fields Virology Fourth Edition* Vol. 1 (eds D. M. Knipe *et al.*) Ch. 46, 1487-1531 (Lippincott Williams & Wilkins, 2001).
- 14 Chanturiya, A. N. *et al.* PB1-F2, an influenza A virus-encoded proapoptotic mitochondrial protein, creates variably sized pores in planar lipid membranes. *Journal of virology* **78**, 6304-6312, doi:10.1128/JVI.78.12.6304-6312.2004 (2004).
- 15 Chen, W. *et al.* A novel influenza A virus mitochondrial protein that induces cell death. *Nat Med* **7**, 1306-1312, doi:10.1038/nm1201-1306 (2001).
- 16 Le, L., Lee, E., Schulten, K. & Truong, T. N. Molecular modeling of swine influenza A/H1N1, Spanish H1N1, and avian H5N1 flu N1 neuraminidases bound to Tamiflu and Relenza. *PLoS Curr* **1**, RRN1015, doi:10.1371/currents.RRN1015 (2009).
- 17 Taubenberger, J. K. *et al.* Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**, 889-893, doi:10.1038/nature04230 (2005).
- 18 Wright, P. F. & Webster, R. G. in *Fields Virology 4th Edition* Vol. 1 (eds D. M. Knipe *et al.*) 1533-1579 (Lippincott Williams & Wilkins, 2001).
- 19 Swayne, D. E. in *Fenner's Veterinary Virology Fourth Edition* (eds N. J. MacLachlan & E. J. Dubovi) Ch. 353-370, (Academic Press, 2011).

- 20 Maines, T. R. *et al.* Multiple genes contribute to the virulent phenotype observed in ferrets of an H5N1 influenza virus isolated from Thailand in 2004. *Virology* **413**, 226-230, doi:10.1016/j.virol.2011.02.005 (2011).
- 21 Zhang, H., Hale, B. G., Xu, K. & Sun, B. Viral and host factors required for avian H5N1 influenza A virus replication in mammalian cells. *Viruses* **5**, 1431-1446, doi:10.3390/v5061431 (2013).
- 22 Wu, X. L. *et al.* Hygrothermal environment may cause influenza pandemics through immune suppression. *Human vaccines & immunotherapeutics* **11**, 2641-2646, doi:10.1080/21645515.2015.1084452 (2015).
- 23 Qiu, J. *et al.* Spatiotemporal pattern and risk factors of the reported novel avian-origin influenza A(H7N9) cases in China. *Preventive veterinary medicine* **115**, 229-237, doi:10.1016/j.prevetmed.2014.03.030 (2014).
- 24 Slingenbergh, J. I., Gilbert, M., de Balogh, K. I. & Wint, W. Ecological sources of zoonotic diseases. *Revue scientifique et technique* **23**, 467-484 (2004).
- 25 Taubenberger, J. K., Reid, A. H., Krafft, A. E., Bijwaard, K. E. & Fanning, T. G. Initial genetic characterization of the 1918 "Spanish" influenza virus. *Science* **275**, 1793-1796 (1997).
- 26 Vidic, J. *et al.* Amyloid Assemblies of Influenza A Virus PB1-F2 Protein Damage Membrane and Induce Cytotoxicity. *The Journal of biological chemistry* **291**, 739-751, doi:10.1074/jbc.M115.652917 (2016).
- 27 Deventhiran, J., Kumar, S. R., Raghunath, S., Leroith, T. & Elankumaran, S. PB1-F2 Protein Does Not Impact the Virulence of Triple-Reassortant H3N2 Swine Influenza Virus in Pigs but Alters Pathogenicity and Transmission in Turkeys. *Journal of virology* **90**, 222-231, doi:10.1128/JVI.01551-15 (2015).
- 28 Buehler, J. *et al.* Influenza A virus PB1-F2 protein expression is regulated in a strain-specific manner by sequences located downstream of the PB1-F2 initiation codon. *Journal of virology* **87**, 10687-10699, doi:10.1128/JVI.01520-13 (2013).
- 29 Vidy, A. *et al.* The Influenza Virus Protein PB1-F2 Increases Viral Pathogenesis through Neutrophil Recruitment and NK Cells Inhibition. *PloS one* **11**, e0165361, doi:10.1371/journal.pone.0165361 (2016).
- 30 Varga, Z. T., Grant, A., Manicassamy, B. & Palese, P. Influenza virus protein PB1-F2 inhibits the induction of type I interferon by binding to MAVS and decreasing mitochondrial membrane potential. *Journal of virology* **86**, 8359-8366, doi:10.1128/JVI.01122-12 (2012).
- 31 Leymarie, O. *et al.* PB1-F2 attenuates virulence of highly pathogenic avian H5N1 influenza virus in chickens. *PloS one* **9**, e100679, doi:10.1371/journal.pone.0100679 (2014).
- 32 Mitzner, D. *et al.* Phosphorylation of the influenza A virus protein PB1-F2 by PKC is crucial for apoptosis promoting functions in monocytes. *Cell Microbiol* **11**, 1502-1516, doi:10.1111/j.1462-5822.2009.01343.x (2009).
- 33 Conenello, G. M., Zamarin, D., Perrone, L. A., Tumpey, T. & Palese, P. A single mutation in the PB1-F2 of H5N1 (HK/97) and 1918 influenza A viruses contributes to increased virulence. *PLoS pathogens* **3**, 1414-1421, doi:10.1371/journal.ppat.0030141 (2007).
- 34 Steinhauer, D. A. Role of hemagglutinin cleavage for the pathogenicity of influenza virus. *Virology* **258**, 1-20, doi:10.1006/viro.1999.9716 (1999).

- 35 Zhang, Y. *et al.* A single amino acid at the hemagglutinin cleavage site contributes to the pathogenicity and neurovirulence of H5N1 influenza virus in mice. *Journal of virology* **86**, 6924-6931, doi:10.1128/JVI.07142-11 (2012).
- 36 Suguitan, A. L., Jr. *et al.* The multibasic cleavage site of the hemagglutinin of highly pathogenic A/Vietnam/1203/2004 (H5N1) avian influenza virus acts as a virulence factor in a host-specific manner in mammals. *Journal of virology* **86**, 2706-2714, doi:10.1128/JVI.05546-11 (2012).
- 37 Horimoto, T. & Kawaoka, Y. Biologic effects of introducing additional basic amino acid residues into the hemagglutinin cleavage site of a virulent avian influenza virus. *Virus research* **50**, 35-40 (1997).
- 38 Bogs, J. *et al.* Highly pathogenic H5N1 influenza viruses carry virulence determinants beyond the polybasic hemagglutinin cleavage site. *PloS one* **5**, e11826, doi:10.1371/journal.pone.0011826 (2010).
- 39 Phung, T. T. *et al.* Key role of regulated upon activation normal T-cell expressed and secreted, nonstructural protein1 and myeloperoxidase in cytokine storm induced by influenza virus PR-8 (A/H1N1) infection in A549 bronchial epithelial cells. *Microbiology and immunology* **55**, 874-884, doi:10.1111/j.1348-0421.2011.00396.x (2011).
- 40 Na-Ek, P. *et al.* The activation of B cells enhances DC-SIGN expression and promotes susceptibility of B cells to HPAI H5N1 infection. *Biochemical and biophysical research communications* **490**, 1301-1306, doi:10.1016/j.bbrc.2017.07.017 (2017).
- 41 Liu, Q., Zhou, Y. H. & Yang, Z. Q. The cytokine storm of severe influenza and development of immunomodulatory therapy. *Cellular & molecular immunology* **13**, 3-10, doi:10.1038/cmi.2015.74 (2016).
- 42 Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution* **30**, 2725-2729, doi:10.1093/molbev/mst197 (2013).
- 43 Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution* **16**, 111-120 (1980).
- 44 Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology* **157**, 105-132 (1982).
- 45 Gasteiger, E. *et al.* in *The Proteonomics Protocols Handbook* (ed J. M. Walker) 571-607 (Humana Press, 2005).
- 46 Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858, doi:10.1038/nprot.2015.053 (2015).
- 47 Reid, A. H., Fanning, T. G., Hultin, J. V. & Taubenberger, J. K. Origin and evolution of the 1918 "Spanish" influenza virus hemagglutinin gene. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 1651-1656 (1999).
- 48 Garten, R. J. *et al.* Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* **325**, 197-201, doi:10.1126/science.1176225 (2009).
- 49 Kimura, M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution* **16**, 111-120 (1980).

Table 1

GenBank sequence data for Spanish flu, pandemic-2009 H1N1, and bird flu H5N1 analyzed in this study

Classified in the text	Strain nomenclature	Segment genome	Accession Number	Reference
Spanish Flu (SF)	(A/Brevig Mission/1/1918(H1N1))	PB2	DQ208309	17
		NS	AF333238	17
		NA	AF250356	17
		MP	AY130766	17
		NP	AY744935	17
		PB1	DQ208310	17
		PA	DQ208311	17
		HA	AF117241	47
Pandemic H1N1 2009 (PDM)	A/South Carolina/1/18 (H1N1)	NP	FJ969512	48
		MP	FJ969513	48
		NS	FJ969514	48
		PA	FJ969515	48
		PB2	FJ969516	48
		NA	FJ969517	48
		HA	GQ117044	48
		PB1	GQ377049	48
Bird Flu H5N1 (BF)	A/goose/Guangdong/1/1996(H5N1)	PB2	AF144300	48
		PB1	AF144301	48
		PA	AF144302	48
		NP	AF144303	48
		NA	AF144304	48
		HA	AF144305	48
		MP	AF144306	48
		NS	AF144307	48

Table 2

Genetic distances, numbers of amino acid differences, and isoelectric points of all structural and nonstructural proteins of Spanish-H1N1, pandemic-2009 H1N1, and bird flu-H5N1 viruses.

Gene	Protein length (residues)			Genetic distance		Number of amino acid difference		Isoelectric point		
	SF	PDM	BF	SF-PDM	SF-BF	SF-PDM	SF-BF	SF	PDM	BF
PB2	759	759	759	0.14	0.18	25	19	9.9	9.8	9.7
PB1	757	757	757	0.21	0.18	28	19	9.5	9.6	9.5
PB1-F2	90	11	90	NA	0.11	27	NA	10.7	NA	11
PA	716	716	716	0.18	0.16	25	24	5.3	5.3	5.4
HA	566	566	568	0.22	0.49	80	195	6	6.9	6.3
NP	498	498	498	0.13	0.14	28	10	9.5	9.6	9.7
NA	469	469	469	0.2	0.19	59	44	5.4	6.1	5.9
M1	252	252	252	0.11	0.1	11	5	9.6	9.8	9.8
M2	97	97	97	0.06	0.06	8	6	4.9	4.5	4.6
NS1	230	219	230	0.17	0.4	37	72	8.2	5.7	6.4
NS2	121	121	121	0.12	0.21	12	20	5	5.1	4.8

Genetic distances were calculated with the Kimura 2-parameter model⁴⁹. Isoelectric points were calculated with Peptide Calculator (<http://www.bachem.com>).

SF: Spanish-H1N1; PDM: pandemic-2009 H1N1; BF: bird flu-H5N1.

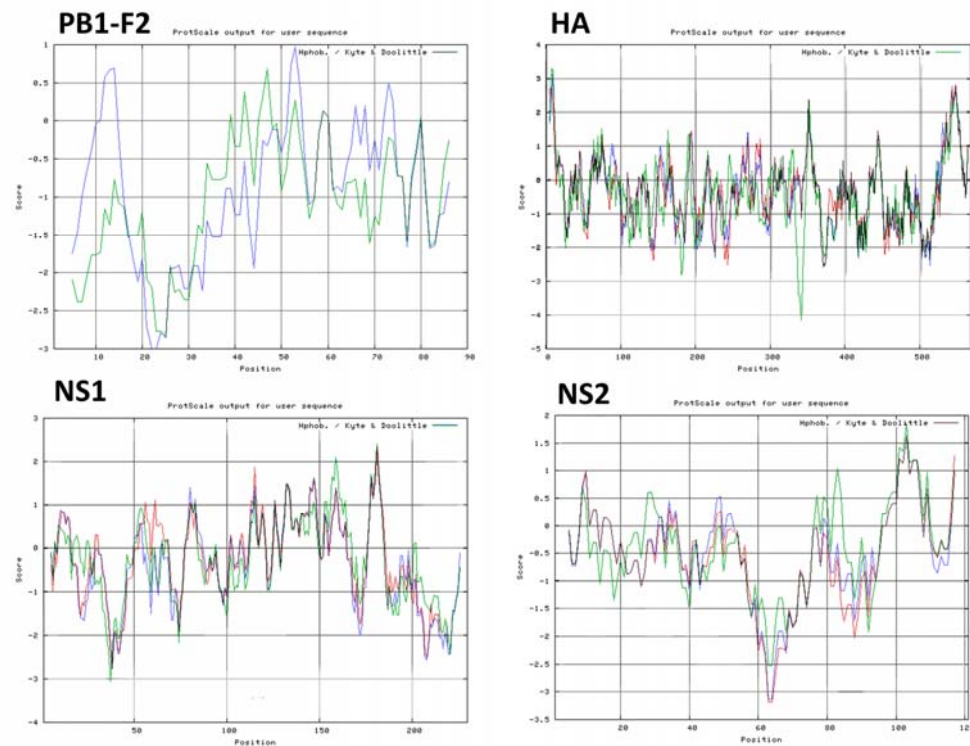


Figure 1

Superimposed hydrophobicity plots of Spanish (blue), pandemic-2009 (red), and bird flu (green) viruses for PB1-F2 (top left), HA (top right), NS1 (bottom left), and NS2 (bottom right). Hydrophobicity plot of each peptide was calculated with the Kyte and Doolittle method ⁴⁴ using the online software ProtScale ⁴⁵.

(<http://web.expasy.org/protscale/>), and the results were superimposed with Adobe Creative Cloud 2015 (Adobe System Inc.).

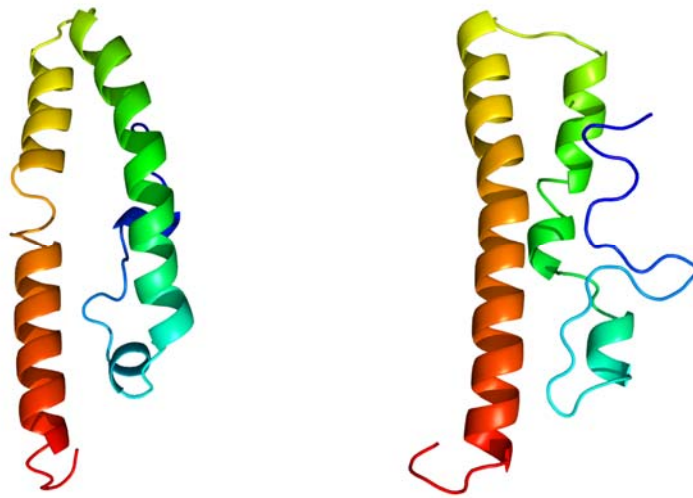


Figure 2

Cartoon peptide modeling of PB1-F2 of SF (left) and BF (right). Images are colored by inverted rainbow from N- to C-terminus. Protein modeling was performed with the online resource PYRE2 (<http://www.sbg.bio.ic.ac.uk>)⁴⁶. Protein models were visualized with RasWin 2.7.5.2 (www.rasmol.org).

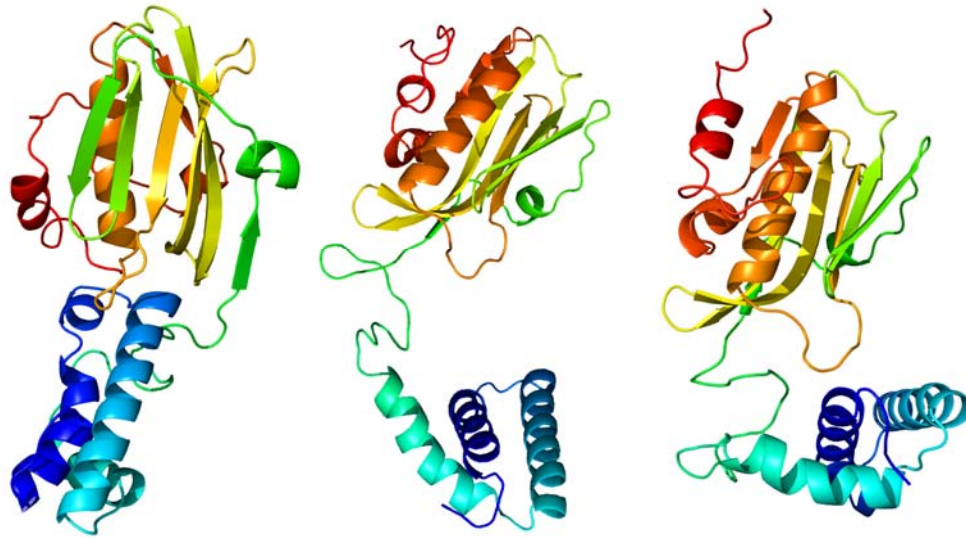


Figure 3

Cartoon peptide modeling of NS1 of SF (left), PDM (middle), and BF (right). Images are colored by inverted rainbow from N- to C-terminus. Protein modeling was performed with the online resource PYRE2 (<http://www.sbg.bio.ic.ac.uk>)⁴⁶. Protein models were visualized with RasWin 2.7.5.2 (www.rasmol.org).

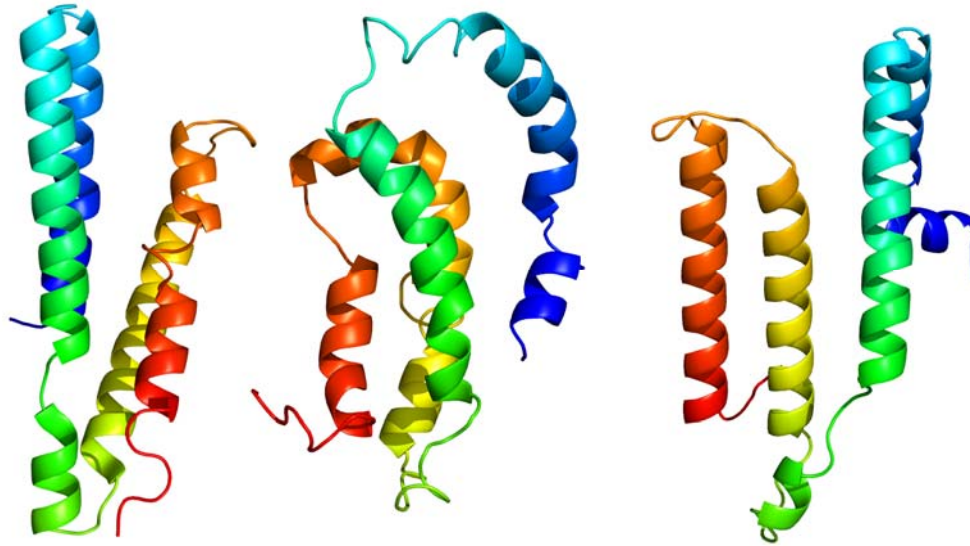


Figure 4

Cartoon peptide modeling of NS2 of SF (left), PDM (middle), and BF (right). Images are colored by inverted rainbow from N- to C-terminus. Protein modeling was performed with the online resource PYRE2 (<http://www.sbg.bio.ic.ac.uk>)⁴⁶. Protein models were visualized with RasWin 2.7.5.2 (www.rasmol.org).

Supplementary figure

Superimposed hydrophobicity plots of Spanish (blue), pandemic-2009 (red), and bird flu (green) viruses for all proteins. Hydrophobicity plot of each peptide was calculated with the Kyte and Doolittle method ⁴⁴ using the online software ProtScale ⁴⁵ (<http://web.expasy.org/protscale/>), and the results were superimposed with Adobe Creative Cloud 2015 (Adobe System Inc.). First row: PB2 (left), PB1 (middle), PB1-F2 (right); second row: PA (left), HA (middle), NP (right); third row: NA (left), M1 (middle), M2 (right); forth row: NS1 (left), NS2 (middle).