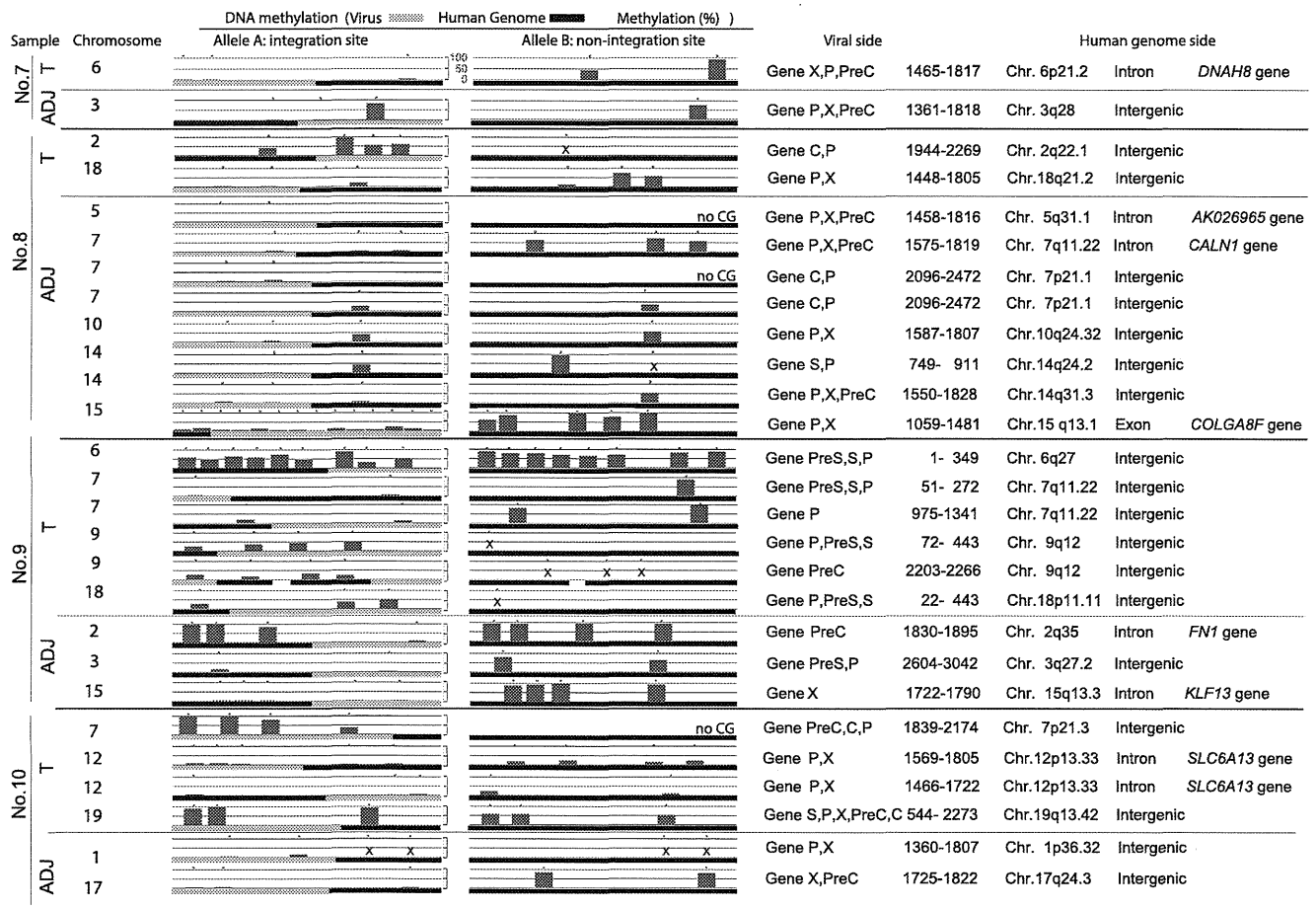


Figure 4. Allele-specific methylation analysis of the tumor (*T*) and adjacent nontumor (ADJ) sample genomes. The methylation levels of the HBV and human genomes for the integrated and unintegrated alleles in four paired tumor and adjacent nontumor samples (sample nos. 7–10) are shown. Detailed results of the HBV integrants (*PreC, Precore; C, Core; PreS, Presurface; S, Surface; X, X*) and flanking host genomes (position, chromosome, location of the genome, and gene names) are shown. The HBV genome became significantly methylated when integrated into highly methylated human genome regions, but not when integrated into unmethylated human genome regions. (×) The desired quantitative methylation levels were not obtained because of technical difficulties with the sequences that were being analyzed.

The dynamic changes in DNA methylation described here have a major functional impact on the biological behavior of HBV and underlie the molecular mechanisms that control infection or enable tumorigenesis. These findings may significantly impact public health given that millions of people worldwide are carriers of HBV. Distinct DNA methylation profiles may exist, for example, between primary HCCs in Japanese patients and those of other nationalities. Additional studies are needed to address this issue, and research into the influence of other environmental factors is required.

Increased viral DNA methylation is present in cancers associated with DNA viruses, including human papilloma virus types 16 and 18 (HPV 16 and 18) (Fernandez et al. 2009; Mirabello et al. 2012), Epstein-Barr virus (Uozaki and Fukayama 2008; Fernandez et al. 2009), and human T-lymphotropic virus 1 (Taniguchi et al. 2005). An analysis of the haplotype-resolved genome and epigenome of the aneuploid HeLa cervical cancer cell line revealed that an amplified, highly rearranged region of chromosome 8q24.21 harboring an integrated HPV18 genome likely represents the tumor-initiating event (Adey et al. 2013). Whether the dynamic changes in DNA methylation observed in cells with integrated HBV genomes also occur in human cells infected by other viruses is an interesting question for further study. We anticipate that our assay will be a powerful tool for this purpose and have successfully detected integrated HPV sequences in the genomes of cervical cancer cell lines (Y Watanabe, H Yamamoto, F Itoh, and N Suzuki, unpubl.).

This study provides novel mechanistic insights into HBV-mediated hepatocarcinogenesis, which may have preventive and therapeutic applications for carriers of HBV and patients with HBV-HCC, as it suggests that epigenetic alterations provide candidate biochemical markers and therapeutic targets. This study, together with a recent global survey of HBV integration events (Ding et al. 2012; Fujimoto et al. 2012; Jiang et al. 2012; Sung et al. 2012; Toh et al. 2013), provides a foundation for the further experimentation and mechanistic understanding of HBV-HCC.

Methods

Cell lines and primary tissues

The PLC/PRF/5 (Alexander) human hepatoma cell line was obtained from the Japanese Collection of Research Bioresources (JCRB). HepG2.2.15 cells, kindly gifted by Professor Stephan Urban

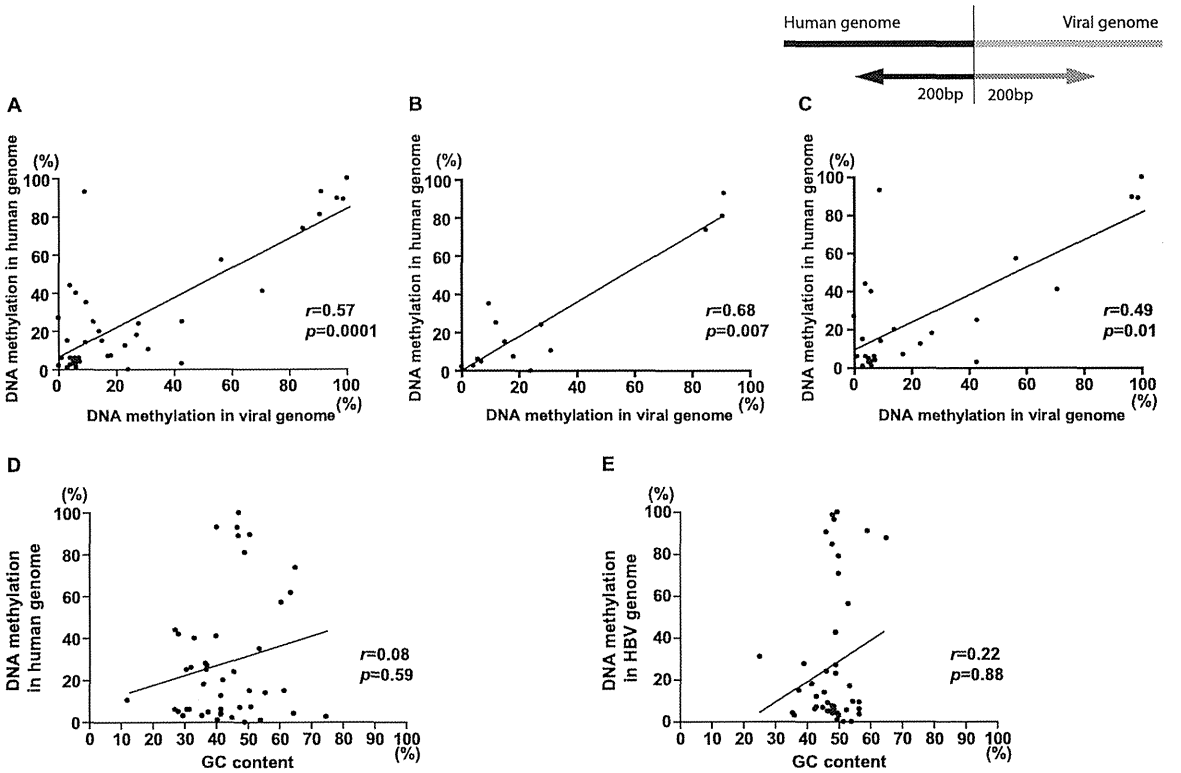


Figure 5. Correlation analysis between the methylation pattern of the integrated HBV DNA and that of the human genome. DNA fragments, including 200 bp of the HBV DNA and 200 bp of the human genome around the boundary, were analyzed for average methylation and GC content. (*A*) A correlation between the average methylation of the HBV DNA and that of the human genome in combined two cell lines and eight clinical samples ($n = 40$, $r = 0.57$, $P = 0.0001$, 95%CI = 0.3091–0.7545). (*B*) A correlation between the average methylation of the HBV DNA and that of the human genome in two cell lines ($n = 14$, $r = 068$, $P = 0.007$, 95%CI = 0.2233–0.8946). (*C*) A correlation between the average methylation of the HBV DNA and that of the human genome in eight clinical samples ($n = 26$, $r = 0.49$, $P = 0.01$, 95%CI = 0.1222–0.7463). (*D*) No correlation between the average methylation and GC contents in the human genome in the combined two cell lines and eight clinical samples ($n = 45$, $r = 0.08$, $P = 0.59$, 95%CI = −0.2253–0.3745). (*E*) No correlation between the average methylation and GC contents in the viral genome in the combined two cell lines and eight clinical samples ($n = 47$, $r = 0.22$, $P = 0.88$, 95%CI = −0.3151–0.2751).

at University Hospital Heidelberg, was derived from HepG2 cells transfected with a plasmid carrying four 5′-3′ tandem copies of the HBV genome (Koike et al. 1994). Cell lines were maintained in appropriate media containing 10% fetal bovine serum in plastic culture plates. Primary tissues from tumor and adjacent tissues were obtained at the time of the clinical procedures. Informed consent was obtained from all the patients before specimen collection. This study was approved by the institutional review board. DNA was extracted using the standard phenol–chloroform method. The concentration and quantity of extracted DNA were measured using a NanoDrop spectrophotometer (NanoDrop Technologies).

MCAM analysis

MCAM analysis was conducted as previously described (Oishi et al. 2012). A detailed protocol of MCA was previously described (Toyota et al. 1999). We used a custom human promoter array (G4426A-02212; Agilent Technologies) comprising 36,579 probes corresponding to 9021 unique genes. The probes on the array were selected to recognize SmaI/XmaI fragments mainly derived from sequences near gene transcription start sites. Five micrograms of genomic DNA was digested with 100 U of methylation-sensitive restriction endonuclease SmaI (New England Biolabs) for 24 h at 25°C, which cleaves unmethylated DNA leaving blunt ends (CCC/GGG). Subsequently, the DNA was digested with 20 U of methylation-insensitive restriction endonuclease XmaI for 6 h at 37°C, creating sticky ends (C/CCGGG). Five hundred milligrams of digested DNA was ligated using 50 μL of RMCA12 (5′-CCGGGCA GAAAG-3′)/RMCA24 (5′-CCACCGCCATCCGAGCCTTTCTGC-3′) primers and T4 DNA ligase (TaKaRa Bio) for 16 h at 16°C. After filling in the overhanging ends of the ligated DNA fragments at 72°C, the DNA was amplified for 5 min at 95°C followed by 25 cycles of 1-min incubation at 95°C and 3-min incubation at 77°C using 100 pmol of RMCA24 primer. MCA products were labeled with Cy5 (red) for DNA from hepatoma samples (both tumor and adjacent normal) and Cy3 (green) for DNA from human blood mixture of three healthy volunteers using a randomly primed Klenow polymerase reaction (Invitrogen) for 3 h at 37°C. Human CpG island arrays (4 × 44 K) were purchased from Agilent Technologies. Microarray protocols, including labeling, hybridization, and post-hybridization washing procedures, are provided at http://www.agilent.com/. Labeled samples were then hybridized to arrays in the presence of human Cot-1 DNA for 24 h at 65°C. After washing, arrays were scanned using an Agilent DNA microarray scanner and analyzed using Agilent Feature Extraction software (FE version 9.5.1.1, Agilent Technologies) at St. Marianna University School of Medicine. We used GeneSpring software (Agilent) for choosing candidate genes after normalization of the raw data.

DNA methylation analysis

Hidden Markov models have been successfully used to partition genomes into segments of comparable stochastic structure (Durbin et al. 1998). Using these models for sequence analysis performed

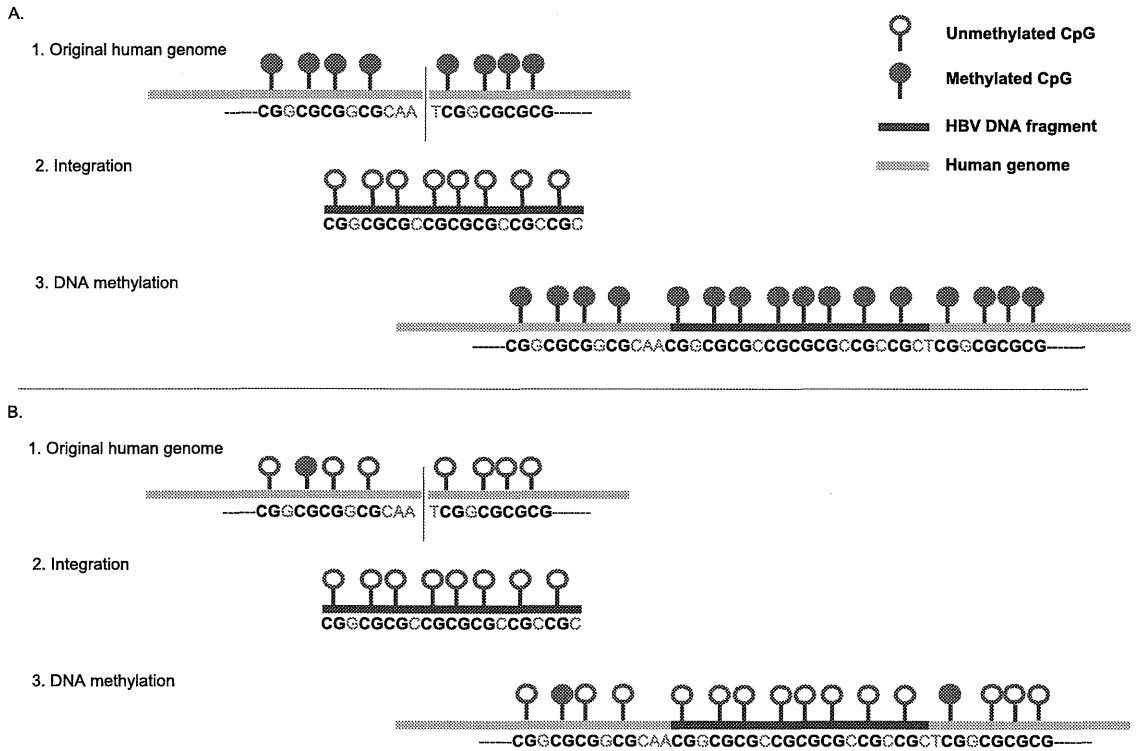


Figure 6. Schema of DNA methylation at HBV integrants and flanking human genomic sequences. (*A*) DNA hypermethylation was seen in both the integrated HBV fragment and human genome (original human genome shows dense methylation). The HBV genome often showed significant methylation when integrated into highly methylated host sites. (*B*) DNA hypermethylation was rarely seen in the integrated HBV fragment and human genome (original human genome shows low methylation). The HBV genome remained largely unmethylated when integrated into unmethylated host sites.

on the CpG plugin of bioinformatics software Geneious 5.5.8 (Biomatters), CpG islands were searched in the HBV genome (Kearse et al. 2012). Bisulfite PCR was performed using an EpiTect Bisulfite Kit (Qiagen) according to the manufacturer's protocol. One microliter of bisulfite-treated DNA was used as a template. The primers used for amplifying CpG sequences in the *HBx* gene are described in Supplemental Table 1. After PCR, the biotinylated strand was captured on streptavidin-coated beads (Amersham Bioscience) and incubated with sequencing primers (Supplemental Table 1). The pyrosequencing reactions were performed using the PyroMark Q24 and/or PyroMark Q24 advanced (Qiagen). Pyrosequencing quantitatively measures the methylation status of several CpG sites in a given sequence. These adjacent sites usually show highly concordant methylation. Therefore, the mean percentage of methylation at detected sites was used as a representative value for each sequence.

LINE1 and *Alu*Yb8 methylation analysis

The LINE1 and *Alu*Yb8 methylation levels, as measured by pyrosequencing, are good indicators of the cellular levels of 5-methylcytosine (i.e., the global DNA methylation level). To quantify relatively high LINE1 and *Alu*Yb8 methylation levels, we used pyrosequencing technology (Igarashi et al. 2010). PCR and subsequent pyrosequencing for LINE1 and *Alu*Yb8 were performed using the PyroMark kit (Qiagen). This assay amplifies a region of the LINE1 or *Alu*Yb8 elements that includes three CpG sites. The PCR was conducted as follows: 45 cycles for 20 sec at 95°C, for 20 sec at 50°C, and for 20 sec at 72°C, followed by 5 min at 72°C. The biotinylated PCR product was purified and converted to single strands to serve as a template for the pyrosequencing reaction using the pyrosequencing vacuum prep tool (Qiagen). The pyrosequencing reactions were performed using the PyroMark Q24 and/or PyroMark Q24 advanced (Qiagen). The percentage of Cs relative to the total sum of the Cs and Ts at each CpG site was calculated. The average of the percentages of Cs at the three CpG sites was used to represent the overall LINE1 and *Alu*Yb8 methylation levels in each sample.

FISH analysis of HBV integration

We developed a FISH analysis method to detect HBV DNA and demonstrate its presence in PLC/PRF/5 cells (Supplemental Fig. 3). The slides were pretreated with hydrogen peroxide and rinsed in 1× phosphate-buffered saline (PBS) to minimize background and quench endogenous peroxidase activity. To remove the excess cytoplasm, the slides were treated with pepsin and then fixed with 1% formaldehyde in PBS/$MgCl_2$. The slides were then dehydrated in an ethanol series (70%, 90%, and 100%) for 3 min at each step. The probes were designed based on the reference HBV sequence in PLC/PRF/5 DNA that is available from the Methylyzer 1.0 website (http://gbrowse.bioinfo.cnio.es/cgi-bin/VIRUS/HBV/). The FISH probes were prepared by combining the PCR-labeled probes (Supplemental Table 1), human Cot-1 DNA, and salmon sperm DNA. The probes were precipitated and mixed with hybridization buffer, and the probe DNA cocktail was denatured for 5 min at 95°C. The DNA on the slides was denatured by soaking in 70% formamide/2× SSC for 3 min at 74°C. The slides were immediately immersed in freshly prepared ice-cold 70% ethanol for 3 min, followed by 3-min immersions in 90% and then 100% ethanol. The denatured probe DNA was applied to the dry denatured slides and covered with a coverslip. The hybridization was performed for 16 h at 37°C.

Tyramide signal amplification (TSA)–FISH

TSA (tyramide signal amplification) detection kits were obtained from PerkinElmer. TSA-FISH detection was performed following the manufacturer's protocols with minor modifications. High stringency washes (0.1× SSC) were used to reduce the background, and TNT buffer (0.1 M Tris-HCl at pH 7.5, 0.15 M NaCl, 0.05% Tween 20) was adjusted to pH 7.0–7.5. The biotin- or DIG-labeled probes were detected using streptavidin-HRP or anti-DIG-HRP in TNB (0.1 M Tris-HCl at pH 7.5, 0.15 M NaCl, 0.05% blocking reagent [supplied in the kit]) for 30 min at room temperature and washed twice for 5 min each in TNT buffer. For the tyramide amplification procedure, the slide was covered with tyramide solution (Tyr-Bio, 1:50) for 10 min at room temperature. The tyramide solution was removed, and the slides were washed twice for 5 min each with TNT at room temperature. Fluorochrome-conjugated streptavidin (stAv-Alexa 488) diluted in TNB was used to detect the Tyr-Bio. The slides were incubated for 30 min at room temperature, washed with TNT buffer twice for 5 min each at room temperature, and covered with an anti-fade reagent containing DAPI (Speel et al. 1997; Schriml et al. 1999).

Identification of the chromosomal locations of viral–host junctions

The viral–host junctions were amplified using primers specific for human *Alu*Yb8 repetitive sequences and HBV X regions (Supplemental Table 1; Minami et al. 1995; Murakami et al. 2004). One microliter of genomic DNA solution served as a template in the subsequent PCR. We used touchdown PCR for most of the assays. All PCR assays included a denaturation step for 30 sec at 95°C, followed by an annealing step at various temperatures for 30 sec and an extension step for 30 sec at 72°C. PCR products were analyzed using electrophoresis through 1% agarose gels. PCR products were ligated to pCR-XL-TOPO vector DNA (TOPO XL PCR Cloning kit; Invitrogen) and transformed into competent cells. Positive colonies were selected and isolated using a QIA prep Spin Miniprep Kit (Qiagen). Direct sequence analysis of TOPO-TA cloning products was performed using a 3130 genetic analyzer (Applied Biosystems) (Watanabe et al. 2011). All sequences were searched for matches with HBV and pCR-XL-TOPO sequences using Geneious 5.5.8 (Biomatters) sequence analysis and assembly software and the BLAST program available on the UCSC Genome Browser (http://genome.ucsc.edu/).

Analysis of HBV DNA integration site sequences using NGS

Agilent's SureSelect target enrichment system is a highly efficient hybrid selection technique for optimizing NGS. We used this system and 12,391 custom baits covering the DNA sequences of HBV genotypes A to J and PLC/PRF/5 HBV sequences and optimized experiments for a GS FLX Titanium system (Roche). PLC/PRF/5, HepG2.2.15, and four paired tumor and nontumor samples (sample nos. 7, 8, 9, and 10 in Fig. 1B) were analyzed.

DNA methylation analysis of the integrated HBV genome as well as the adjacent human genome

DNA methylation was analyzed using bisulfite pyrosequencing (Oishi et al. 2012). The pyrosequencing reactions were performed using the PyroMark Q24 and/or PyroMark Q24 advanced (Qiagen). PLC/PRF/5, HepG2.2.15, and four paired tumor and nontumor samples (sample nos. 7, 8, 9, and 10 in Fig. 1B) were analyzed. Primers for methylation analysis of integration sites in PLC/PRF/5 are shown in Supplemental Table 1.

DNA methylation analysis of orthologous loci

Methylation levels of orthologous loci in HepG2.2.15 cells and in PBLs of a healthy volunteer at the same (empty) target sites of PLC/PRF/5 cells were analyzed using bisulfite pyrosequencing. Similarly, methylation levels of orthologous loci in PLC/PRF/5 cells and in PBLs at the same (empty) target sites of HepG2.2.15 cells were analyzed.

Allele-specific DNA methylation analysis of the integrated HBV genome as well as the adjacent human genome

Allele-specific DNA methylation was analyzed as described previously (Yamada and Ito 2011). The pyrosequencing reactions were performed using the PyroMark Q24 and/or PyroMark Q24 advanced (Qiagen).

Correlation analysis between the methylation pattern of the integrated HBV DNA and that of the human genome

DNA fragments, including 200 bp of the integrated HBV DNA and 200 bp of the human genome around the boundary, were analyzed for average methylation, GC content, and repetitive sequences in cell lines and clinical samples. RepeatMasker was used to identify repetitive elements in genomic sequences (AFA Smit, R Hubley, P Green, unpubl.). The Spearman correlation coefficient was used to assess correlations between the average methylation of the HBV DNA and that of the human genome. Correlations between GC content or repetitive sequences in the HBV DNA and the human genome were analyzed by using the Spearman correlation coefficient for continuous variables, and $P < 0.05$ was considered significant. All statistical analyses were performed using PRISM software for Windows, version 4 (GraphPad Prism).

Chromatin structure at the integrated HBV site

Using Bander software (Cheung et al. 2001; Furey and Haussler 2003), we analyzed the chromatin structure at the integrated HBV site in PLC/PRF/5 and HepG2.2.15.

Data access

All raw sequence data from this study have been submitted to the DDBJ Japanese Genotype-phenotype Archive (JGA; http://trace.ddbj.nig.ac.jp/jga/index_e.html) under accession number JGAS00000000015. Array data have been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE59405.

Acknowledgments

We thank Drs. K. Watashi and T. Wakita for cell lines and Drs. T. Takayama, K. Takasaki, and S. Kawasaki for clinical samples. We also thank Drs. N. Matsumoto and N. Yamada-Ohkawa, as well as other members of the laboratory, for advice and suggestions. A part of the data used for this research was originally obtained by a research project of Hiroyuki Yamamoto and Yoshiyuki Watanabe led by Professor Fumio Itoh and is available at the website of the NBDC/JST (http://biosciencedbc.jp/en/). This work was supported in part by the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (JSPS KAKENHI grant no. 23590964 to H. Yotsuyanagi).

Author contributions: Y.W. conceived the study, designed and performed the experiments, analyzed the data, and wrote the manuscript. H. Yamamoto designed the experiments, analyzed the data and wrote the manuscript. R.O. performed the experiments

and analyzed the data. M.T. provided intellectual support. M.Y., N.K., S.T., and A.S. provided clinical samples. H. Yotsuyanagi designed the experiments, analyzed the data and wrote the manuscript. K.K. provided intellectual support. F.I. supervised all aspects of the study.

References

Adey A, Burton JN, Kitzman JO, Hiatt JB, Lewis AP, Martin BK, Qiu R, Lee C, Shendure J. 2013. The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line. *Nature* **500:** 207–211.

Burgers WA, Blanchon L, Pradhan S, de Launoit Y, Kouzarides T, Fuks F. 2007. Viral oncoproteins target the DNA methyltransferases. *Oncogene* **26:** 1650–1655.

Cheung VG, Nowak N, Jang W, Kirsch IR, Zhao S, Chen XN, Furey TS, Kim UJ, Kuo WL, Olivier M, et al. 2001. Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature* **409:** 953–958.

Ding D, Lou X, Hua D, Yu W, Li L, Wang J, Gao F, Zhao N, Ren G, Li L, et al. 2012. Recurrent targeted genes of hepatitis B virus in the liver cancer genomes identified by a next-generation sequencing-based approach. *PLoS Genet* **8:** e1003065.

Doerfler W. 2008. In pursuit of the first recognized epigenetic signal-DNA methylation: a 1976 to 2008 synopsis. *Epigenetics* **3:** 125–133.

Doerfler W, Remus R, Müller K, Heller H, Hohlweg U, Schubbert R. 2001. The fate of foreign DNA in mammalian cells and organisms. *Dev Biol* **106:** 89–97.

Durbin R, Eddy S, Krogh A, Mitchison G. 1998. *Biological sequence analysis*. Cambridge University Press, Cambridge, UK.

Fernandez AF, Rosales C, Lopez-Nieva P, Graña O, Ballestar E, Ropero S, Espada J, Melo SA, Lujambio A, Fraga MF, et al. 2009. The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer. *Genome Res* **19:** 438–451.

Fujimoto A, Totoki Y, Abe T, Boroevich KA, Hosoda F, Nguyen HH, Aoki M, Hosono N, Kubo M, Miya F, et al. 2012. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat Genet* **44:** 760–764.

Furey TS, Haussler D. 2003. Integration of the cytogenetic map with the draft human genome sequence. *Hum Mol Genet* **12:** 1037–1044.

Gatza ML, Chandhasin C, Ducu RI, Marriott SJ. 2005. Impact of transforming viruses on cellular mutagenesis, genome stability, and cellular transformation. *Environ Mol Mutagen* **45:** 304–325.

Hilleman MR. 2004. Strategies and mechanisms for host and pathogen survival in acute and persistent viral infections. *Proc Natl Acad Sci* **101:** 14560–14566.

Igarashi S, Suzuki H, Niinuma T, Shimizu H, Nojima M, Iwaki H, Nobuoka T, Nishida T, Miyazaki Y, Takamaru H, et al. 2010. A novel correlation between LINE-1 hypomethylation and the malignancy of gastrointestinal stromal tumors. *Clin Cancer Res* **16:** 5114–5123.

Jiang Z, Jhunjhunwala S, Liu J, Haverty PM, Kennemer MI, Guan Y, Lee W, Carnevali P, Stinson J, Johnson S, et al. 2012. The effects of hepatitis B virus integration into the genomes of hepatocellular carcinoma patients. *Genome Res* **22:** 593–601.

Kan Z, Zheng H, Liu X, Li S, Barber TD, Gong Z, Gao H, Hao K, Willard MD, Xu J, et al. 2013. Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma. *Genome Res* **23:** 1422–1433.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28:** 1647–1649.

Kim CM, Koike K, Saito I, Miyamura T, Jay DG. 1991. HBx gene of hepatitis B virus induces liver cancer in transgenic mice. *Nature* **351:** 317–320.

Koike K, Moriya K, Iino S, Yotsuyanagi H, Endo Y, Miyamura T, Kurokawa K. 1994. High-level expression of hepatitis B virus HBx gene and hepatocarcinogenesis in transgenic mice. *Hepatology* **19:** 810–819.

Lau C-C, Sun T, Ching AK, He M, Li JW, Wong AM, Co NN, Chan AW, Li PS, Lung RW, et al. 2014. Viral-human chimeric transcript predisposes risk to liver cancer development and progression. *Cancer Cell* **25:** 335–349.

Li S, Mao M. 2013. Next generation sequencing reveals genetic landscape of hepatocellular carcinomas. *Cancer Lett* **340:** 247–253.

Lupberger J, Hildt E. 2007. Hepatitis B virus–induced oncogenesis. *World J Gastroenterol* **13:** 74–81.

Minami M, Poussin K, Brechot C, Paterlini P. 1995. A novel PCR technique using *Alu*-specific primers to identify unknown flanking sequences from the human genome. *Genomics* **29:** 403–408.

Mirabello L, Sun C, Ghosh A, Rodriguez AC, Schiffman M, Wentzensen N, Hildesheim A, Herrero R, Wacholder S, Lorincz A, et al. 2012. Methylation of human papillomavirus type 16 genome and risk of cervical precancer in a Costa Rican population. *J Natl Cancer Inst* **104:** 556–565.

Murakami Y, Minami M, Daimon Y, Okanoue T. 2004. Hepatitis B virus DNA in liver, serum, and peripheral blood mononuclear cells after the clearance of serum hepatitis B virus surface antigen. *J Med Virol* **72:** 203–214.

Nakagawa H, Shibata T. 2013. Comprehensive genome sequencing of the liver cancer genome. *Cancer Lett* **340:** 234–240.

Oishi Y, Watanabe Y, Yoshida Y, Sato Y, Hiraishi T, Oikawa R, Maehata T, Suzuki H, Toyota M, Niwa H, et al. 2012. Hypermethylation of Sox17 gene is useful as a molecular diagnostic application in early gastric cancer. *Tumour Biol* **33:** 383–393.

Orend G, Kuhlmann I, Doerfler W. 1991. Spreading of DNA methylation across integrated foreign (adenovirus type 12) genomes in mammalian cells. *J Virol* **65:** 4301–4308.

Schriml LM, Padilla-Nash HM, Coleman A, Moen P, Nash WG, Menninger J, Jones G, Ried T, Dean M. 1999. Tyramide signal amplification (TSA)-FISH applied to mapping PCR-labeled probes less than 1 kb in size. *Biotechniques* **27:** 608–613.

Sells MA, Chen ML, Acs G. 1987. Production of hepatitis B virus particles in Hep G2 cells transfected with cloned hepatitis B virus DNA. *Proc Natl Acad Sci* **84:** 1005–1009.

Speel EJ, Ramaekers FC, Hopman AH. 1997. Sensitive multicolor fluorescence in situ hybridization using catalyzed reporter deposition (CARD) amplification. *J Histochem Cytochem* **45:** 1439–1446.

Sung WK, Zheng H, Li S, Chen R, Liu X, Li Y, Lee NP, Lee WH, Ariyaratne PN, Tennakoon C, et al. 2012. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nat Genet* **44:** 765–769.

Taniguchi Y, Nosaka K, Yasunaga J, Maeda M, Mueller N, Okayama A, Matsuoka M. 2005. Silencing of human T-cell leukemia virus type I gene transcription by epigenetic mechanisms. *Retrovirology* **2:** 64.

Tao Q, Robertson KD. 2003. Stealth technology: how Epstein-Barr virus utilizes DNA methylation to cloak itself from immune detection. *Clin Immunol* **109:** 53–63.

Toh ST, Jin Y, Liu L, Wang J, Babrzadeh F, Gharizadeh B, Ronaghi M, Toh HC, Chow PK, Chung AY, et al. 2013. Deep sequencing of the hepatitis B virus in hepatocellular carcinoma patients reveals enriched integration events, structural alterations and sequence variations. *Carcinogenesis* **34:** 787–798.

Toyota M, Ho C, Ahuja N, Jair KW, Li Q, Ohe-Toyota M, Baylin SB, Issa JP. 1999. Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Res* **59:** 2307–2312.

Uozaki H, Fukayama M. 2008. Epstein-Barr virus and gastric carcinoma-viral carcinogenesis through epigenetic mechanisms. *Int J Clin Exp Pathol* **1:** 198–216.

Watanabe Y, Castoro RJ, Kim HS, North B, Oikawa R, Hiraishi T, Ahmed SS, Chung W, Cho MY, Toyota M, et al. 2011. Frequent alteration of MLL3 frameshift mutations in microsatellite deficient colorectal cancer. *PLoS ONE* **6:** e23320.

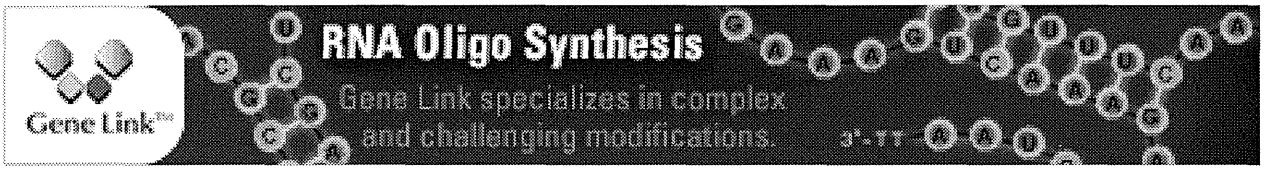
Yamada Y, Ito T. 2011. Highly efficient PCR assay to discriminate allelic DNA methylation status using whole genome amplification. *BMC Res Notes* **4:** 179.

Received March 8, 2014; accepted in revised form December 29, 2014.

DNA methylation at hepatitis B viral integrants is associated with methylation at flanking human genomic sequences

Yoshiyuki Watanabe, Hiroyuki Yamamoto, Ritsuko Oikawa, et al.

Genome Res. published online February 4, 2015
Access the most recent version at doi:10.1101/gr.175240.114



CellPress

Cell Host & Microbe
Article

Circulating Avian Influenza Viruses Closely Related to the 1918 Virus Have Pandemic Potential

Tokiko Watanabe,[1,2,12] Gongxun Zhong,[1,12] Colin A. Russell,[3,12] Noriko Nakajima,[4] Masato Hatta,[1] Anthony Hanson,[1] Ryan McBride,[5] David F. Burke,[6] Kenta Takahashi,[4] Satoshi Fukuyama,[2] Yuriko Tomita,[2] Eileen A. Maher,[1] Shinji Watanabe,[2,7] Masaki Imai,[8,9] Gabriele Neumann,[1] Hideki Hasegawa,[4] James C. Paulson,[5] Derek J. Smith,[6] and Yoshihiro Kawaoka[1,2,10,11,*]

[1]Department of Pathobiological Sciences, School of Veterinary Medicine, University of Wisconsin-Madison, 575 Science Drive, Madison, WI 53711, USA
[2]ERATO Infection-Induced Host Responses Project, Japan Science and Technology Agency, Saitama 332-0012, Japan
[3]Department of Veterinary Medicine, University of Cambridge, Madingley Road, Cambridge CB3 0ES, UK
[4]Department of Pathology, National Institute of Infectious Diseases, Shinjuku, Tokyo 162-8640, Japan
[5]Department of Cell and Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA
[6]Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK
[7]Laboratory of Veterinary Microbiology, Department of Veterinary Sciences, University of Miyazaki, Miyazaki 889-2192, Japan
[8]Department of Veterinary Medicine, Faculty of Agriculture, Iwate University, Iwate 020-8550, Japan
[9]Influenza Virus Research Center, National Institute of Infectious Diseases, Tokyo 208-0011, Japan
[10]Division of Virology, Department of Microbiology and Immunology, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan
[11]Department of Special Pathogens, International Research Center for Infectious Diseases, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108-8639, Japan
[12]Co-first authors
*Correspondence: kawaokay@svm.vetmed.wisc.edu
http://dx.doi.org/10.1016/j.chom.2014.05.006

SUMMARY

Wild birds harbor a large gene pool of influenza A viruses that have the potential to cause influenza pandemics. Foreseeing and understanding this potential is important for effective surveillance. Our phylogenetic and geographic analyses revealed the global prevalence of avian influenza virus genes whose proteins differ only a few amino acids from the 1918 pandemic influenza virus, suggesting that 1918-like pandemic viruses may emerge in the future. To assess this risk, we generated and characterized a virus composed of avian influenza viral segments with high homology to the 1918 virus. This virus exhibited pathogenicity in mice and ferrets higher than that in an authentic avian influenza virus. Further, acquisition of seven amino acid substitutions in the viral polymerases and the hemagglutinin surface glycoprotein conferred respiratory droplet transmission to the 1918-like avian virus in ferrets, demonstrating that contemporary avian influenza viruses with 1918 virus-like proteins may have pandemic potential.

INTRODUCTION

Despite having conquered many infectious diseases, we continue to face a threat from novel, previously unrecognized infectious diseases. Most of these so-called emerging infectious diseases are caused by zoonotic pathogens that originate in animals (Jones et al., 2008; Taylor et al., 2001). These pathogens are responsible for various human diseases, such as AIDS, SARS, and pandemic influenza. Animal-origin pathogens must overcome an immense hurdle to cause a zoonosis (i.e., interspecies transmission from natural or intermediate hosts to humans). If such pathogens acquire the ability to transmit among humans, they become an appreciable threat to mankind.

The 1918 influenza pandemic, the most devastating disease outbreak recorded, killed an estimated 40–50 million people worldwide (Johnson and Mueller, 2002; Taubenberger and Morens, 2006). Sequence analyses identified the 1918 pandemic strain as an H1N1 influenza A virus of avian origin, although this conclusion remains controversial (Rabadan et al., 2006; Reid et al., 2004; Smith et al., 2009). Since avian species harbor a large influenza virus gene pool that may contain influenza viral segments encoding proteins with high homology to the 1918 viral proteins (designated as 1918 virus-like virus proteins), the possibility exists for a 1918 virus-like avian virus to emerge in the human population as a pandemic virus; however, the likelihood of such an event remains unknown. To assess the risk of emergence of pandemic influenza viruses reminiscent of the 1918 influenza virus, we examined the properties of influenza viruses composed of avian influenza viral segments that encode proteins with high homology to the 1918 viral proteins (designated as 1918-like avian viruses), which we generated by using reverse genetics.

RESULTS

Generation of a 1918-like Avian Virus Possessing Avian Influenza Viral Segments that Encode Proteins with High Homology to the 1918 Viral Proteins

We first determined whether influenza A virus gene segments that encode proteins with high homology to the 1918 viral proteins exist in the avian influenza virus gene pool. We focused on amino acid sequence comparisons because our goal was to identify avian influenza viral proteins that closely resemble 1918 virus proteins in structure and function and may therefore allow the emergence of a 1918-like virus. To this end, we performed a sequence similarity search using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify the closest relatives to the 1918 viral proteins. Interestingly, for most viral proteins (except for hemagglutinin [HA], neuraminidase [NA], and PB1-F2), we found avian influenza virus proteins that differed from their 1918 counterparts by only a limited number of amino acids (Table S1). For example, for polymerase basis 2 (PB2) we found one avian influenza PB2 protein that differed from 1918 PB2 by eight amino acids. Similarly, we found avian influenza PB1, polymerase acid (PA), nucleoprotein (NP), matrix protein 1 (M1), M2, nonstructural protein 1 (NS1), and NS2 proteins that closely resembled their 1918 counterparts. Most of the viruses from which these proteins were derived were isolated recently, suggesting that 95 years after the devastating 1918 pandemic, avian influenza virus genes encoding 1918-like proteins continue to circulate in nature. For the 1918 HA and NA proteins, we identified the closest avian H1 and N1 relatives, respectively. These proteins differed from the 1918 proteins by 33 and 31–33 amino acids, respectively, a finding that was expected due to the higher evolutionary rates in these genes relative to the other influenza viral genes.

To assess the risk of emergence of a 1918-like virus and to delineate the amino acid changes that are needed for such a virus to become transmissible via respiratory droplets in mammals, we attempted to generate an influenza virus composed of avian influenza viral segments that encoded proteins with high homology to the 1918 viral proteins. In particular, we selected the following genes to generate a 1918-like avian influenza virus (referred to as 1918-like avian virus) (see Supplemental Experimental Procedures): the PB2 segment of A/blue-winged teal/Ohio/926/2002 (H3N8), the PB1 segment of A/blue-winged teal/Alberta (ALB)/286/77 (H3N6), the PA segment of A/pintail duck/ALB/219/77 (H1N1), the NP segment of A/blue-winged teal/Ohio/908/2002 (H1N1), the M segment of A/duck/Germany/113/95 (H9N2), the NS segment of A/canvasback duck/Alberta/102/76 (H3N6), the HA segment of A/pintail duck/ALB/238/79 (H1N1), and the NA segment of A/mallard/duck/ALB/46/77 (H1N1), as described in the Supplemental Information. The resulting virus differs by 8 (PB2), 6 (PB1), 20 (PB1-F2), 9 (PA), 7 (NP), 33 (HA), 31 (NA), 1 (M1), 5 (M2), 4 (NS1), and 0 (NS2) amino acids from the 1918 virus.

The 1918-like avian virus possessing these eight 1918-like avian viral segments was successfully recovered from 293T cells transfected with the plasmids required to generate this virus. The virus grew well in Madin-Darby canine kidney (MDCK) cells and embryonated chicken eggs (8.3 ± 0.1 and 8.6 ± 0.2 $\log_{10}$ plaque-forming units (pfu)/ml at 24 hr postinfection, respectively), and its growth was comparable to that of the 1918 virus (8.8 ± 0.1 and 8.4 ± 0.2 $\log_{10}$ pfu/ml in MDCK cells and eggs, respectively, at 24 hr postinfection).

The 1918-like Avian Virus Exhibits Intermediate Pathogenicity Compared with the 1918 Virus and an Authentic Avian Virus in Mouse and Ferret Models

To assess the pathogenicity of the 1918-like avian virus, we first determined the mouse lethal dose 50 (MLD_{50}; the dose required to kill 50% of infected mice) values of the 1918-like avian influenza virus and authentic 1918 virus. As a control for the authentic avian influenza virus, we used A/duck/ALB/35/76 (H1N1; DK/ALB) because this avian strain was well characterized in a previous study (Van Hoeven et al., 2009). The 1918-like avian influenza virus showed intermediate pathogenicity (with an MLD_{50} of 5.5 $\log_{10}$ pfu) compared with DK/ALB (with an MLD_{50} of 6.8 $\log_{10}$ pfu) and the 1918 virus (with an MLD_{50} of 2.7 $\log_{10}$ pfu). These data indicate that the avian influenza virus genes that were selected because of their close relationship with 1918 virus proteins not only cooperate at a functional level, but also support pathogenicity higher than that of an authentic avian influenza virus.

The ferret is considered the best current model for influenza virus infection because infected animals exhibit symptoms that resemble those of humans infected with influenza A virus. Therefore, we next tested the pathogenicity of the 1918-like avian virus in ferrets. Ferrets (three animals per group) were intranasally inoculated with 10^6 pfu/500 μl of the 1918-like avian virus, 1918 virus, or DK/ALB virus (Figure S1, available online). All animals infected with the 1918 virus became symptomatic; their body weights declined drastically (the mean maximum body weight loss was 17.5% ± 2.2%), and one of them died on day 8 postinfection (Figure S1A). By contrast, none of the ferrets infected with DK/ALB exhibited noticeable clinical signs (no appreciable body weight loss in any of the animals; Figure S1C), whereas animals infected with the 1918-like avian virus became symptomatic and showed substantial body weight loss (the mean maximum body weight loss was 11.9% ± 3.9%; see Table 1 and Figure S1B). Of the three viruses tested, the 1918 virus replicated the most efficiently in the upper and lower respiratory tract ($p < 0.01$; Figure 1 and Table S2). The 1918- and 1918-like avian virus-infected ferrets displayed numerous viral antigen-positive cells in tracheal, bronchial, bronchiolar, and glandular epithelial cells and necrotized changes in some trachea-bronchial glands on day 6 postinfection (Figure 2A), whereas the 1918-like avian virus-infected ferrets presented few antigen-positive cells on day 3 postinfection (Figure S2). In the lungs, no significant differences in pathologic changes were detected between the 1918 and 1918-like avian virus groups (Figures 2A, 2B, and S2). In contrast, the DK/ALB infection caused only minimal to mild bronchitis or bronchiolitis with little viral antigen expression on days 3 and 6 postinfection (Figures 2A, 2B, and S2). Taken together with the results of recent studies that showed no appreciable clinical signs in ferrets infected with more commonplace avian influenza viruses of various subtypes (Koçer et al., 2012; Watanabe et al., 2013), the finding that the 1918-like avian influenza virus is of intermediate pathogenicity in mammals may suggest that the progenitor of the 1918 virus was an unusual avian influenza virus whose

Table 1. Pathogenicity and Transmissibility of Reassortants and Mutants of 1918-like Avian Virus in Ferrets

	Virus-Inoculated Ferrets		Contact Ferrets	
Virus	Weight Loss (%)[a]	Survival (No. of Dead/Total)	Virus Detection in Nasal Wash (Positive and Total Numbers)	Seroconversion (Positive and Total Numbers)[b]
1918	17.5 ± 2.2	1/3	2 of 3	2 of 3
1918-like Avian[c]	11.9 ± 3.9	0/3	0 of 3	0 of 3
DK/ALB	1.9 ± 3.3	0/3	0 of 3	0 of 3
1918 PB2/Avian	17.8 ± 2.2	0/3	0 of 3	0 of 3
1918 HA/Avian	16.9 ± 5.8	2/3	0 of 3	0 of 3
1918 PB2:HA/Avian	18.7 ± 4.0	2/3	0 of 3	0 of 3
1918 PB2:HA:NA/Avian	18.6 ± 3.7	0/3	2 of 3	2 of 3
1918(3P+NP):HA/Avian	21.2 ± 4.4	2/3	1 of 3	1 of 3
1918-like Avian PB2-627K	10.3 ± 3.6	0/3	0 of 3	0 of 3
1918-like Avian PB2-627K HA-89ED/190D/225D	19.4 ± 9.2	1/3	1 of 3	1 of 3
1918-like Avian PB2-627K/684D:HA-89ED/113SN/ 190D/225D/265DV:PA-253M	21.0 ± 5.1	0/3	2 of 3	2 of 3

See also Figures S1 and S3 and Table S3. For each pair of ferrets, one animal was intranasally inoculated with 10^6 pfu of virus (0.5 ml) (virus-inoculated ferret), and 1 day later a naive ferret was placed in an adjacent cage (contact ferret). Virus-inoculated ferrets were monitored for 14 days to determine survival rates and body weight changes.

[a]Maximum percentage weight loss (mean ± SD) is shown.

[b]Serum samples were collected 18 days after infection; HI assays were carried out with homologous virus and turkey red blood cells.

[c]The 1918-like avian virus possessing the avian influenza viral segments that encoded proteins with high homology to the 1918 viral proteins described in the Supplemental Experimental Procedures was generated by using reverse genetics. The effectiveness of current vaccines and antiviral drugs against the 1918-like avian virus was examined (see Tables S4 and S5).

pathogenicity in mammals was higher than that of most avian influenza viruses.

The 1918 PB2 and HA Genes Contribute to Enhanced Pathogenicity and Transmissibility in Ferrets

To identify the 1918 viral segments that are responsible for the intermediate pathogenicity of the 1918-like avian virus relative to the 1918 and authentic avian viruses, we examined the pathogenicity of reassortant viruses possessing 1918 viral genes in the genetic background of the 1918-like avian virus in ferrets. We focused on the HA and PB2 genes, which are known to play important roles in the adaptation of avian influenza viruses to mammals (Hatta et al., 2001; Li et al., 2005; Matrosovich et al., 2006; Neumann and Kawaoka, 2006; Subbarao et al., 1993). We generated 1918 PB2/Avian, 1918 HA/Avian, and 1918 PB2:HA/Avian viruses, which possess the PB2, HA, or PB2 and HA genes of the 1918 virus, respectively, and the remaining genes from the 1918-like avian virus (Figure S3). We also created 1918 PB2:HA:NA/Avian and 1918(3P+NP):HA/Avian viruses, which possess the 1918 virus PB2, HA, and NA genes, or the 1918 virus PA, PB1, PB2, NP, and HA genes, respectively, and the remaining genes from the 1918-like avian virus (Figure S3). All animals infected with these viruses became symptomatic; they lost appetite and body weight (Table 1 and Figure S1). Two of three ferrets died upon infection with the 1918 HA/Avian, 1918 PB2:HA/Avian, and 1918(3P+NP):HA/Avian viruses (Table 1). In the ferret tracheae, the mean virus titers of all reassortant viruses possessing the 1918 PB2 gene or the 1918 HA gene were higher than that of the 1918-like avian virus on day 3 postinfection (Table S2), although these differences were not statistically significant. These results suggest that the 1918 PB2 and HA genes confer high pathogenicity to the 1918-like avian virus in the ferret model.

For an influenza virus to cause a pandemic, it must achieve efficient human-to-human transmission. Therefore, we tested transmissibility in ferrets of the authentic 1918 virus, as well as the 1918-like avian virus and its reassortants. Three animals were each intranasally inoculated with 10^6 pfu of virus. On day 1 after infection, a naive ferret was housed in a cage adjacent to each of the infected ferrets. This setup prevented direct contact between animals but allowed the spread of influenza virus through respiratory droplets. Viral titers were determined in nasal washes collected from both the inoculated and contact ferrets every other day postinfection/contact (for up to 9 days). The 1918 virus was recovered from two of the three contact ferrets (Figure 3A), demonstrating respiratory droplet transmission of the 1918 virus. In contrast, the 1918-like avian and DK/ALB viruses failed to transmit among ferrets; no virus was detected in nasal washes collected from contact animals, although the inoculated animals did shed virus (Figures 3B and 3C). Similarly, no virus was recovered from contact animals for the 1918 PB2/Avian, 1918 HA/Avian, and 1918 PB2:HA/Avian virus groups (Figures S4A–S4C). However, virus was recovered from one of the three contact animals in the 1918(3P+NP):HA/Avian and two of three contact ferrets in the 1918 PB2:HA:NA/Avian virus group (Figures S4D and S4E), suggesting potential roles for the

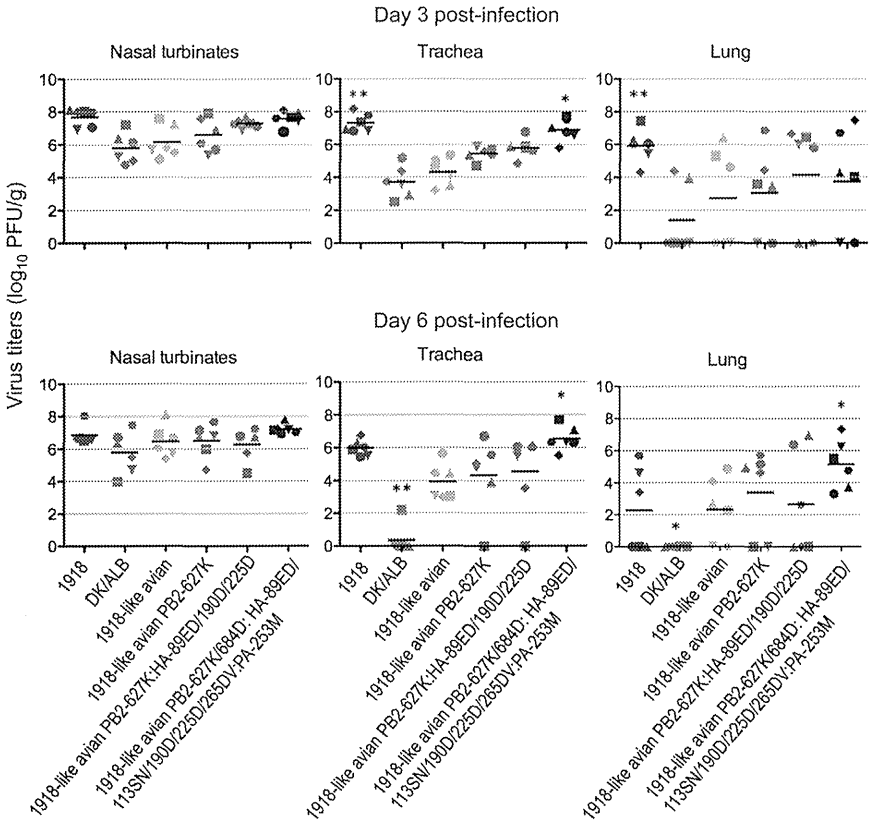


Figure 1. Replicative Ability of 1918 Virus, an Authentic Avian Virus, 1918-like Avian Virus, and 1918-like Avian Mutant Viruses

Virus replication in respiratory organs of ferrets. Ferrets were intranasally infected with 10^6 pfu of virus. Six animals per group were euthanized on days 3 and 6 postinfection for virus titration. Virus titers in nasal turbinates, tracheae, and lungs were determined by plaque assay in MDCK cells. Horizontal bars indicate the mean virus titers. Asterisks indicate virus titers significantly different from those of the 1918-like avian virus (*$p < 0.05$; **$p < 0.01$). See also Table S2.

RNA replication complex, HA, and NA in virus transmission among ferrets. Taken together, our data suggest that the 1918 PB2 and HA genes confer enhanced pathogenicity and transmissibility to the 1918-like avian virus in ferrets.

Identification of Amino Acid Substitutions Potentially Associated with the Efficient Transmission of the 1918-like Avian Virus in Ferrets

HA and PB2 are known to play important roles in restricting viral transmission from avian species to humans (Van Hoeven et al., 2009). The receptor-binding specificity of the HA protein is a major determinant of influenza viral host range (Matrosovich et al., 2008). E-to-D and G-to-D mutations at positions 190 and 225 of HA (H3 numbering), respectively, are essential for avian virus HAs of the H1 subtype to bind to human-type receptors (Matrosovich et al., 2000). In addition, Lys at position 627 of the PB2 protein is important for avian viruses to efficiently replicate in mammalian cells and at the lower temperatures of the upper human airway (Hatta et al., 2001, 2007; Subbarao et al., 1993). To test whether these mammalian-adapting mutations in HA and PB2 affect the replicative ability, pathogenicity, and transmissibility of the 1918-like avian virus, we attempted to generate mutant 1918-like reassortant avian viruses possessing these amino acid substitutions (i.e., 1918-like avian PB2-627K and 1918-like avian PB2-627K:HA-190D/225D viruses). 1918-like avian PB2-627K virus was generated upon inoculation of the supernatant of transfected 293T cells into embryonated chicken eggs (Figure S3). We were unable to generate 1918-like avian PB2-627K:HA-190D/225D virus in embryonated chicken eggs, but we were able to obtain this virus after propagation of 293T cell supernatant in MDCK cells. However, this virus consisted of a mixed virus population possessing E or D at position 89 of HA (89ED) in addition to the PB2-627K and HA-190D/225D mutations (designated as 1918-like avian PB2-627K:HA-89ED/190D/225D; Figure S3 and Table 2). These mutant viruses replicated well in the nasal turbinates and tracheal tissues of ferrets (Figure 1 and Table S2). Infection of ferrets with 1918-like avian PB2-627K and 1918-like avian PB2-627K:HA-89ED/190D/225D caused substantial body weight loss (10.3% ± 3.6% and 19.4% ± 9.2%, respectively; Table 1) and appreciable pathologic changes in the trachea and lungs of the infected animals (Figures 2A and S2). We tested the transmissibility of these viruses and found that no virus was recovered from contact ferrets for the 1918-like avian PB2-627K virus group. Interestingly, for the 1918-like avian PB2-627K:HA-89ED/190D/225D virus group, virus and seroconversion were detected for one of the three contact ferrets (Table 1, Figures 3E and S4F, and Table S3), indicating its partial transmissibility in this animal model.

We sequenced the virus isolated on days 5 and 9 postcontact from the nasal washes of the contact ferret in the 1918-like avian PB2-627K:HA-89ED/190D/225D virus group and found three additional mutations, HA-S113N, PB2-A684D, and PA-V253M (Table 2). After propagating this virus in MDCK cells, the virus stock possessed the following mutations: PB2-627K, PB2-684D, PA-253M, HA-190D, HA-225D, HA-89E/D, HA-113S/N, and HA-265D/V (the mixed populations of amino acids were found at positions 89, 113, and 265) (designated as 1918-like avian PB2-627K/684D:HA-89ED/113SN/190D/225D/265DV:PA-253M) (Table 2 and Figure S3).

The 1918-like avian PB2-627K/684D:HA-89ED/113SN/190D/225D /265DV:PA-253M virus replicated in trachea and lungs more efficiently than the 1918-like avian virus ($p < 0.05$; Figure 1 and Table S2) and caused severe weight loss in the infected ferrets (maximum body weight loss was 21.0% ± 5.1%), although no infected animals died (Table 1). The 1918-like avian PB2-627K/684D:HA-89ED/113SN/190D/225D/265DV:PA-253M infection caused more progressive inflammation in the lungs of ferrets on

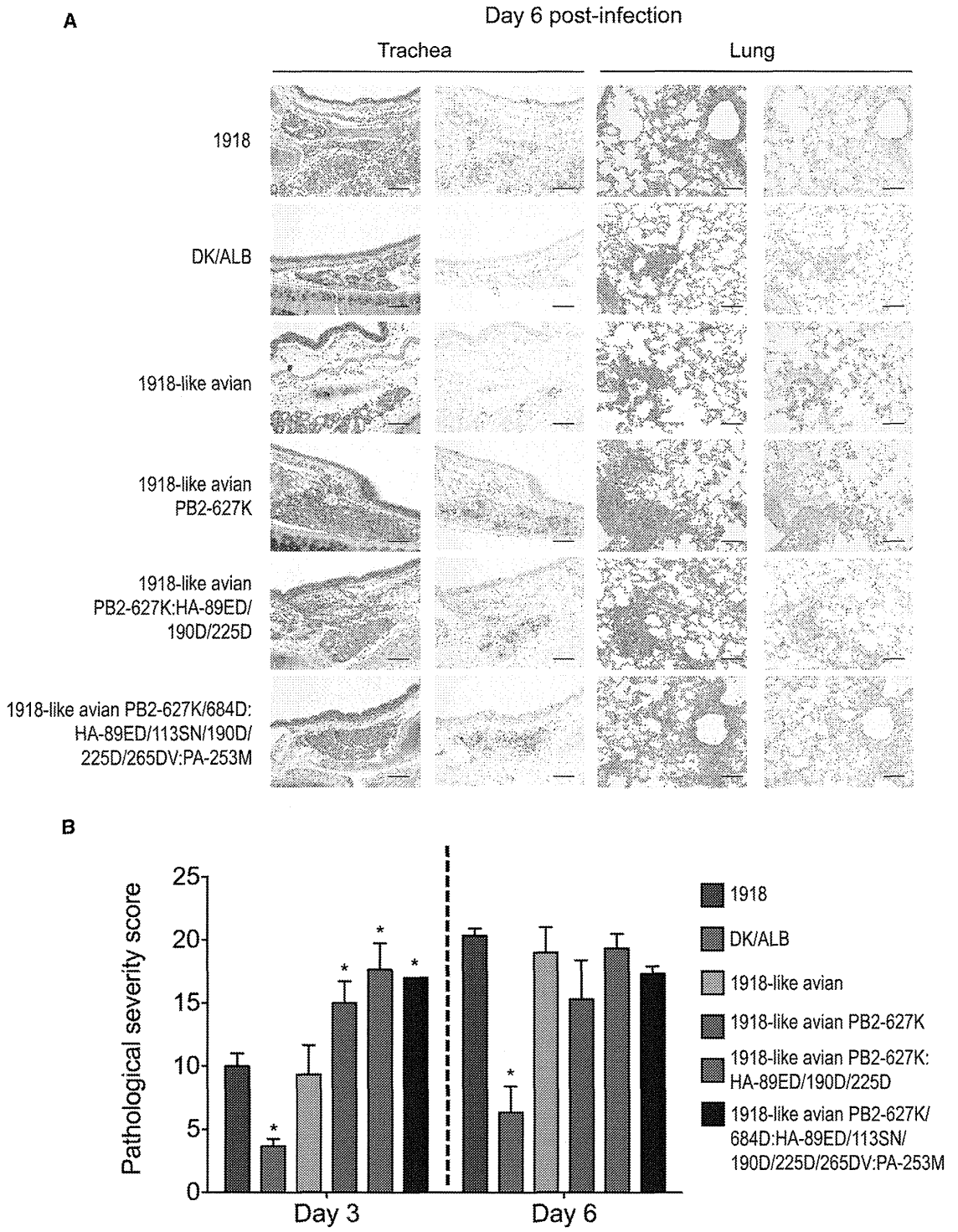


Figure 2. Pathological Analyses of 1918 Virus, an Authentic Avian Virus, 1918-like Avian Virus, and 1918-like Avian Mutant Viruses

(A) Histopathological findings in virus-infected ferrets. Shown are representative pathological changes in tracheae and lungs of ferrets infected with 10^6 pfu of the indicated viruses on day 6 postinfection. Three ferrets per group were infected intranasally with 10^6 pfu of virus, and tissues were collected on day 6 after infection for pathological examination. No virus was detected from the lungs of the DK/ALB-infected ferrets. Left: H&E staining. Right: immunohistochemical staining for influenza viral antigen (NP). Scale bars, 100 μm.

(B) Pathological severity scores for infected ferrets. To represent comprehensive histological changes, respiratory tissue slides were evaluated by scoring pathological changes as described in the Supplemental Experimental Procedures. The sum of the pathologic scores for all five lung lobes was calculated for each ferret. The means ± SD from three ferrets are shown. Asterisks indicate virus pathological scores significantly different from that of the 1918-like avian virus (Dunnett's test; $p < 0.05$). See also Figure S2.

day 3 postinfection compared with that caused by the 1918-like avian virus (Dunnett's test; $p < 0.05$; Figures 2B and S2). The ferrets infected with this mutant virus presented numerous viral antigen-positive cells in tracheal, bronchial, bronchiolar, and glandular epithelial cells on days 3 and 6 postinfections (Figures 2A and S2). We then tested the transmissibility of this virus and found that two of the three contact ferrets were positive for virus between days 5 and 9 after contact; these animals were also seropositive (Figures 3F and S4F and Table S3). Sequence analysis showed that the virus recovered from the contact animal for pair 1 possessed an additional I-to-T amino acid substitution at position 187 of HA (Table 2), which is located at the receptor-binding site of HA (Figure 4A). For the pair 3 contact ferret, the recovered virus possessed an additional T-to-I mutation at position 232 of NP (Table 2). Taken together, our results demonstrate that ten amino acid substitutions (E627K and A684D in PB2; E89D, S113N, I187T, E190D, G225D, and D265V in HA; V253M in PA; and T232I in NP) may be associated with efficient 1918-like avian virus transmission in ferrets.

The Effects of Amino Acid Changes in the HA of the Transmissible 1918-like Avian Viruses on Receptor-Binding Specificity and HA Stability

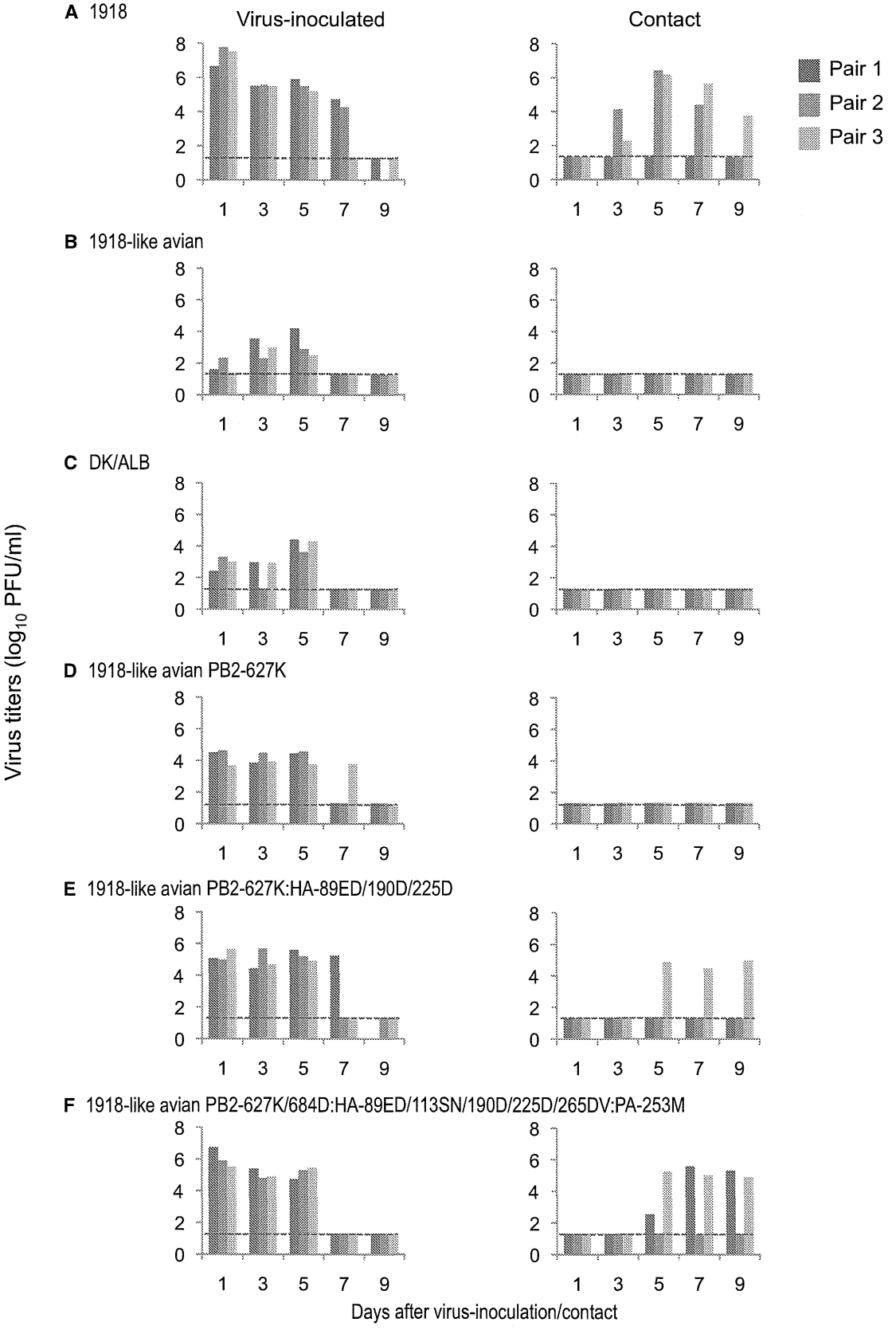
To better understand the molecular mechanisms behind the enhanced replicative ability and transmissibility of 1918-like avian virus carrying HA and PB2 with the amino acid substitutions found in the transmissible 1918-like avian viruses, we examined the effects of these amino acid changes on the functions of HA and the viral polymerases. First, we conducted a glycan microarray to examine the receptor specificity of the 1918-like avian virus HA possessing human-type amino acid substitutions. For this experiment, we used reassortant viruses possessing HA segments derived from 1918, DK/ALB, 1918-like avian HA-190D/225D, 1918-like avian HA-89D/190D/225D, or 1918-like avian HA-89D/113N/190D/225D virus, their NA segment from 1918-like avian virus, and their remaining genes from A/Puerto Rico/8/34 (H1N1; PR8); however, the 1918-like avian virus HA and NA genes were tested in the background of 1918-like avian virus since these genes could not be rescued in the background of PR8. The 1918 HA bound to α2,6-linked sialosides (human-type receptor) (Figure 4B), whereas the 1918-like avian (Figure 4C) and DK/ALB (Figure S5A) HAs bound to α2,3-linked sialosides (avian-type receptor). As reported previously (Chutinimitkul et al., 2010; Liu et al., 2010; Matrosovich et al., 2000; Watanabe et al., 2011; Yang et al., 2010), Gly residues at positions 190 and 225 of HA can confer binding to avian-type receptors, whereas Asp residues at these positions confer binding to human-type receptors. Indeed, the 1918-like avian mutant HAs possessing Asp at positions 190 and 225 efficiently bound to α2,6Gal-sialylated glycans (human-type receptor) (Figure 4D). In addition, mutant HA possessing the additional mutations found in the virus recovered from the contact ferret infected with transmitted 1918-like avian PB2-627K:HA-89ED/190D/225D virus (i.e., HA-89D and HA-113N) also bound to the human-type receptor (Figures 4E and S5B). The 1918-like avian mutant HAs exhibited remarkably similar glycan-binding specificity. For HAs possessing the E-to-D change at position 89 (i.e., HA-89D/113N/190D/225D and HA-89D/190D/225D), there was weak binding to glycan 51 that was not detected with HA-190D/225D (Figures 4 and S5); however, the biological significance is currently unknown.

In addition to receptor binding, another critical function of HA is membrane fusion. A recent study showed a difference in the pH at which HA is activated for fusion between transmissible and nontransmissible viruses (de Vries et al., 2014; Imai et al., 2012). Therefore, we next examined HA fusion activity under different pH conditions by testing the polykaryon formation efficiency of HeLa cells expressing 1918 HA, the 1918-like avian HA, the 1918-like avian mutant HAs, or DK/ALB HA (Figure S5C). The 1918-like avian HA induced membrane fusion at pH 5.3 or below, whereas membrane fusion was observed at pH 5.6 or below in the cells expressing 1918 or DK/ALB HA (Figure S5C). None of the amino acid changes at positions 89, 113, 190, or 225 affected the fusion pH of the 1918-like avian HA (i.e., the value at which HA was activated was pH 5.3) (Figure S5C), suggesting that the amino acid changes found in the HA of the 1918-like avian transmissible mutant have no effect on the pH at which fusion of the 1918-like avian HA is activated.

Our previous studies demonstrated that H5 transmissible mutants exhibit considerable tolerance to high temperature compared with a nontransmissible H5 mutant, suggesting a role for HA stabilization in efficient replication and transmission in ferrets (de Vries et al., 2014; Imai et al., 2012). Therefore, we next tested whether the identified amino acid changes in HA affect its heat stability. For this experiment, we used the same set of recombinant viruses that were used in the glycan array described earlier. We tested the loss of hemagglutination activity and infectivity after incubation at 50°C for various times (Figures S5D and S5E). Introduction of the 190D/225D mutation into 1918-like HA resulted in faster loss of hemagglutination activity, whereas the additional introduction of the HA-89D or HA-89D/113N mutation reversed this trend (Figure S5D). The HA-190D/225D virus lost its infectivity rapidly after a 180 min incubation at 50°C (6.8 $\log_{10}$ decrease in virus titers; Figure S5E), whereas the HA-89D/190D/225D and the HA-89D/113N/190D/225D viruses were more tolerant of high temperature (3.7 and 4.2 $\log_{10}$ decreases in virus titers under the same conditions, respectively; Figure S5E). These results suggest that mutations critical for human-type receptor binding (i.e., HA-190D/225D) reduced HA stability, which was restored by additional mutations associated with respiratory droplet transmissibility (i.e., HA-89D and HA-89D/113N). This trend (i.e., loss of HA stability due to mutations conferring human virus receptor recognition and recovery of HA stability by the acquisition of an additional mutation) is consistent with our previous findings (de Vries et al., 2014; Imai et al., 2012). However, the HA heat stability of 1918 HA, as evaluated by hemagglutination activity, was similar to that of the HA-190D/225D virus (Figure S5D), indicating complex interplay among receptor binding specificity, HA stability, and virus transmissibility in ferrets.

The Effects of Amino Acid Changes in the Viral RNA Polymerase Complex of the Transmissible 1918-like Avian Viruses on Polymerase Activity

The viral polymerase complex affects viral replicative ability and pathogenicity (de Vries et al., 2014; Gabriel et al., 2008; Hatta et al., 2001; Li et al., 2005; Subbarao et al., 1993). To determine



(legend on next page)

Table 2. Amino Acid Changes in Viral Proteins Acquired during the Transmission of 1918-like Avian Viruses in Ferrets

		Amino Acid Sequence and Position									
		HA[a]						PB2		PA	NP
Virus		89	113	187	190	225	265	627	684	253	232
1918[b]		E	S	T	D	D	G	K	A	V	T
1918-like Avian[c]		E	S	I	E	G	D	E	A	V	T
1918-like Avian PB2- 627K:HA-89ED/190D/225D[c]		E/D[d]	S	I	D	D	D	K	A	V	T
1918-like Avian PB2-627K/684D:HA-89ED/113SN/190D/225D:PA253M[c]		E/D[d]	S/N[d]	I	D	D	D	K	D	M	T
1918-like Avian PB2-627K/684D:HA89ED/113SN/190D/225D/265DV:PA253M[f]		E/D[d]	S/N[d]	I	D	D	D/V[d]	K	D	M	T
Pair 1	inoculated	E/D[d]	S/N[d]	I/T[d]	D	D	D/V[d]	K	D	M	T
	contact	E/D[d]	S	T	D	D	V	K	D	M	T
Pair 3	inoculated	E/D[d]	S/N[d]	I	D	D	D/V[d]	K	D	M	T
	contact	E/D[d]	N	I	D	D	D	K	D	M	I

See also Figures S3 and S4.

[a]Amino acid positions of HA are based on H3 HA numbering.

[b]The stock of this egg-grown virus was sequenced.

[c]The stock of this MDCK-grown virus was sequenced.

[d]Indicates a mixed population.

[e]Virus isolated from the nasal washes of the contact ferret in this 1918-like avian PB2-627K:HA-89ED/190D/225D virus group was sequenced.

[f]Virus isolated from the nasal washes of the contact ferret in this 1918-like avian PB2-627K:HA-89ED/190D/225D virus group was propagated in MDCK cells, and the resulting virus stock was sequenced.

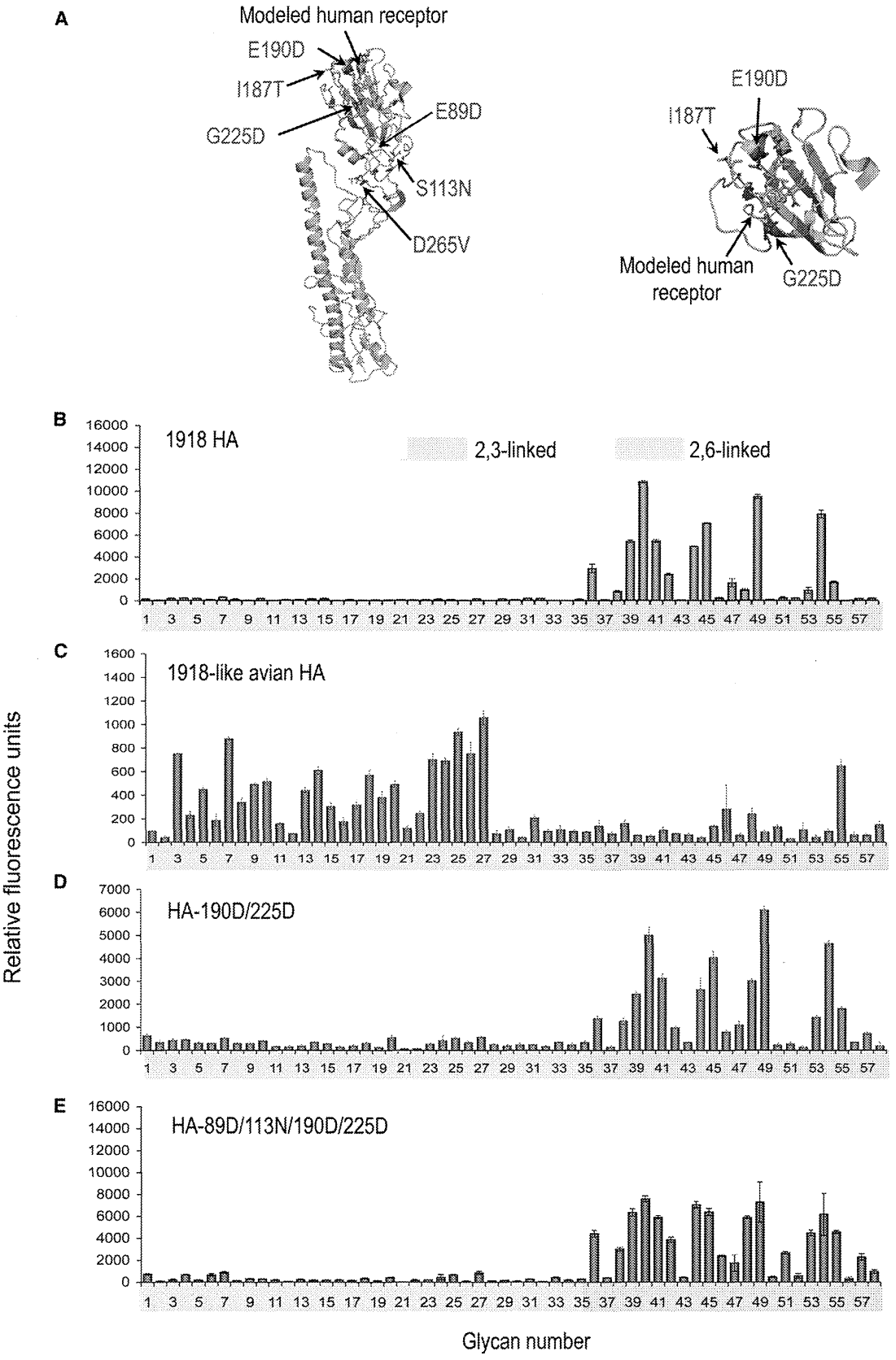
whether amino acid changes in PA and/or PB2 affect the viral polymerase activity of the 1918-like avian virus, we conducted a luciferase activity-based minireplicon assay in human 293T cells as described previously (Ozawa et al., 2007). The polymerase activity of the 1918 polymerase complex was significantly greater than that of the 1918-like avian polymerase complex at 33°C and 37°C (Figure S5F; $p < 0.05$). The PB2-627K mutation strongly increased the polymerase activity of the 1918-like avian polymerase complex ($p < 0.05$), whereas neither PB2-684D nor PA-253M appreciably affected the polymerase activity (Figure S5F). These findings suggest that only the PB2-627K mutation contributes to the enhanced polymerase activity.

The Effectiveness of Current Vaccines and Antiviral Drugs against the 1918-like Avian Virus

From a biosafety perspective, it is important to know whether current control measures (vaccines and antiviral drugs) are effective against the 1918-like avian virus. We therefore examined the reactivity of sera from humans vaccinated with the current influenza vaccine, which includes A/California/04/2009 (H1N1), against the 1918-like avian virus and its mutants. The HI test results revealed that these sera reacted poorly with 1918-like avian virus (i.e., to the same extent as the negative control virus A/mallard/Gurjev/263/82 [H14N5], which possess an HA that is antigenically distant from that of the current vaccine strains) (Table S4). In contrast, the sera reacted robustly with the 1918 virus and the mutant 1918-like avian viruses that possessed additional mutations associated with respiratory droplet transmissibility (i.e., the 1918-like avian PB2-627K:HA-89ED/190D/225D and 1918-like avian PB2-627K/684D:HA-89ED/113SN/190D/225D/265DV:PA-253M viruses), although the HI titers against the 1918-related viruses were substantially lower than those against the homologous strain (i.e., A/California/04/2009) (Table S4). Given that the amino acids at positions 190 and 225 are located close to the antigenic sites of HA, the amino acid substitutions at these positions may alter the antigenicity of the 1918-like avian virus. We next examined the oseltamivir sensitivity of our test viruses. A previous study showed that 1918 virus was highly susceptible to the NA inhibitor oseltamivir (Tumpey et al., 2002). The IC_{50} (half maximal inhibitory concentration) value of the 1918-like avian virus (<10 nM) was similar to that of the 1918 virus and that of A/Kawasaki/UT-K25/2008 (H1N1; an oseltamivir-sensitive control; Table S5). In contrast, the IC_{50} values of the oseltamivir-resistant controls, A/Kawasaki/IMS22B-955/2003 (H3N2) and A/Osaka/180/2009 (H1N1), were 34,600 and 1,360 nM, respectively (Table S5). Because all of the viruses tested in this study possess the NA gene from either the 1918 virus or the 1918-like avian virus, these viruses should be highly susceptible to oseltamivir, so appropriate control measures would be available to combat the 1918-like avian virus used in this study.

Figure 3. Respiratory Droplet Transmission in Ferrets

Groups of three ferrets were infected intranasally with 10^6 pfu of 1918 (A), 1918-like avian (B), DK/ALB (C), 1918-like avian PB2-627K (D), 1918-like avian PB2-627K:HA-89ED/190D/225D (E), and 1918-like avian PB2-627K/684D:HA-89ED/113N/190D/225D/265DV:PA-253M (F) viruses. After 1 day, a naive ferret (contact ferret) was placed in a cage adjacent to each infected ferret. Nasal washes were collected from infected ferrets on day 1 after inoculation and from contact ferrets on day 1 after cohousing and then every other day (for up to 9 days) for virus titration. The lower limit of detection is indicated by the horizontal dashed line. See also Figure S4 and Table S3.



(legend on next page)

Global Distribution and Evolution of 1918-like Avian Viruses

Wild birds harboring a large pool of influenza A virus genomes facilitate the global spread of viruses. Here, we experimentally demonstrated that an avian influenza virus encoding proteins with high homology to the 1918 viral proteins and possessing a limited number of additional mutations exhibited relatively high pathogenicity in mammals and transmissibility between ferrets (Figure 3 and Table 1), suggesting its potential to cause a pandemic. To assess the risk of the emergence of such avian influenza viruses, we examined the prevalence of avian influenza viruses encoding viral proteins that differed from the 1918 virus proteins by a few amino acids. We constructed phylogenetic trees, with a time series of virus isolation (from 1990 to 2011), of amino acid sequences for each viral protein by using the 1918 viral protein as the root sequence (Figures 5A and S6). To assess the spatial patterns and divergence of the 1918-like avian viruses, we also generated geographic maps and histograms to show the distribution of the amino acid differences from their respective 1918 viral proteins (Figures 5B, 5C, and S6). The geographic patterns of amino acid sequence similarity by gene segment showed that avian viruses encoding PB2, PB1, NP, M, and NS genes of closest similarity to those of the 1918-like avian virus have predominantly circulated in North America and Europe, with mostly sporadic circulation in other parts of the world (Figures 5 and S6).

Further, we examined whether there are any avian influenza viruses possessing human-type amino acids, such as PB2-627K, HA-190D, or HA-225D. We found that 168 of 4,293 avian PB2 genes (~4%) possess the PB2-627K mutation and that 142 of those 168 are from H5N1 viruses that were isolated mainly from wild and domestic birds in Europe, the Middle East, and Africa; the others are H1N1 viruses isolated from poultry in the US (Table S7). The HA-190D mutation was found in 9 of 266 avian H1 HA sequences (about 3% of avian H1 HA genes), one that also possessed HA-225D (Table S7). Our BLAST search revealed that all of the avian viruses possessing either HA-190D or HA-225D or both probably originated from swine or human viruses that infected an avian species (Table S7).

Taken together, our results reveal the global distribution of avian influenza virus genes encoding proteins that differ from the 1918 virus proteins by only a few amino acids and the existence of avian influenza viruses possessing human-type amino acid residues (i.e., PB2 E627K, HA E190D, and HA G225D). Considering the fact that many of these avian influenza viruses were isolated in recent years (i.e., from 1990 to 2011), our study demonstrates the continued circulation of avian influenza viruses that possess 1918 virus-like proteins and may acquire 1918 virus-like properties.

DISCUSSION

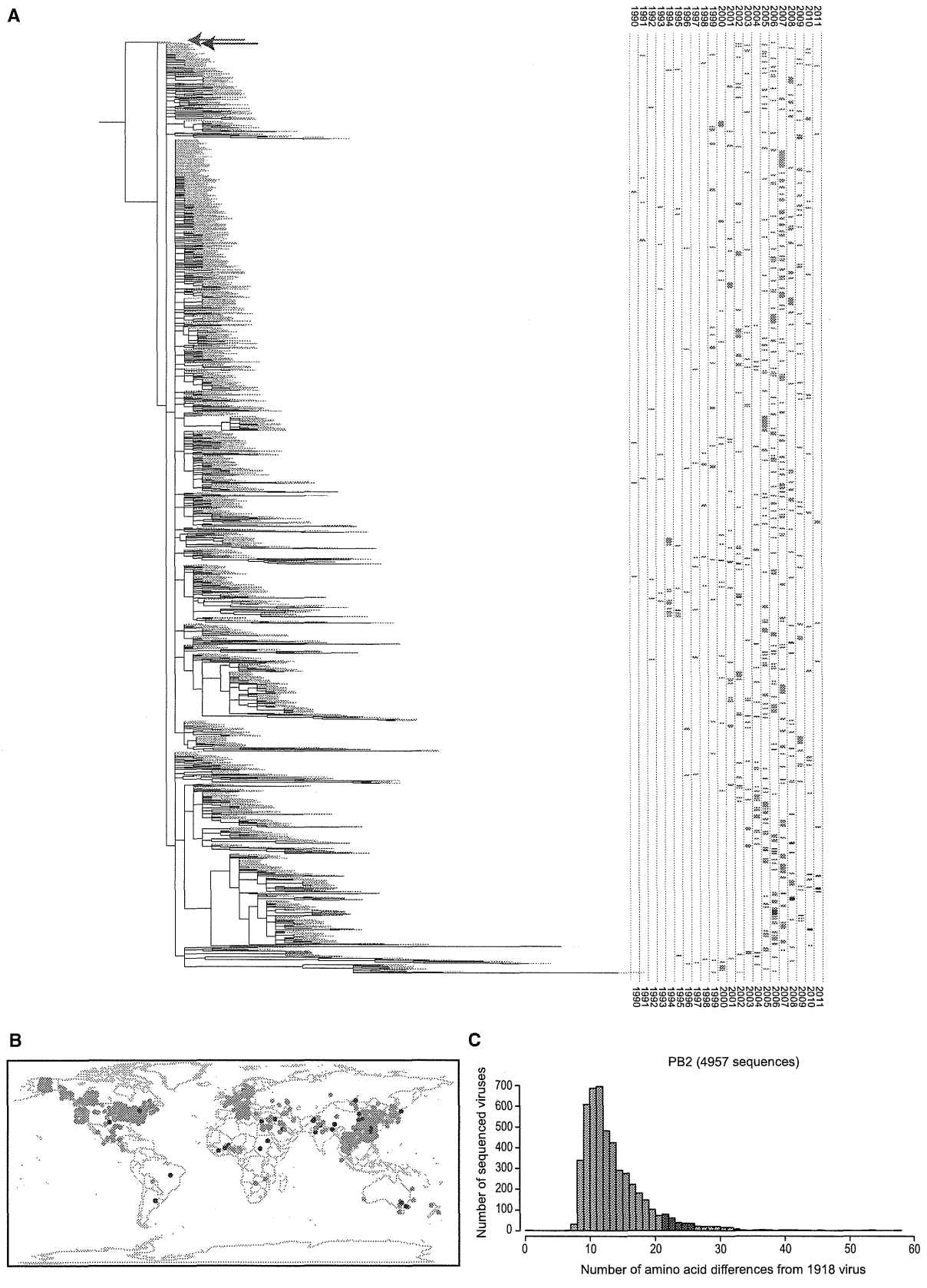
Here, we conducted experiments to assess the risk represented by avian influenza viral genes encoding proteins that closely resemble 1918 virus proteins that still circulate in the avian influenza viral gene pool. We found that a 1918-like avian virus exhibited pathogenicity in mice and ferrets higher than that of an authentic avian virus. Moreover, we demonstrated that acquisition of only a few amino acid substitutions can confer respiratory droplet transmission to 1918-like avian influenza viruses in a ferret model, suggesting that the potential exists for a 1918-like pandemic virus to emerge at any time from the avian virus gene pool.

Generally, in experimental infections, avian influenza viruses replicate poorly in humans, and vice versa (Beare and Webster, 1991; Hinshaw et al., 1984; Murphy et al., 1984; Webster et al., 1978), because host restriction constrains interspecies transmission. Recently, however, a newly emerged H7N9 avian influenza virus appeared to break this host barrier and directly infect humans in China, resulting in a total of 238 confirmed cases and 57 fatalities as of January 24, 2014 (Schnirring, 2014). We and others demonstrated that H7N9 viruses isolated from Chinese patients possess human-adaptive mutations in PB2 and HA, which possibly confer efficient replication of avian influenza viruses in the upper respiratory tract of humans, and caused limited respiratory droplet transmission in ferrets (Belser et al., 2013; Richard et al., 2013; Watanabe et al., 2013; Zhang et al., 2013; Zhu et al., 2013), raising concerns about the pandemic potential of the H7N9 virus. Similarly, other avian influenza viruses, including the 1918-like avian viruses described in this study, have the potential to cause a pandemic if they acquire such human-adaptive mutations. The worst-case scenario is the emergence of a novel avian influenza virus that exhibits high pathogenicity in human, like highly pathogenic avian H5N1 viruses, and efficient transmissibility in humans, like seasonal influenza viruses. To prepare for such a scenario, it is important to understand the molecular mechanisms of pathogenicity and transmissibility of avian influenza viruses. Such information provides support for pandemic preparedness activities (vaccines and antivirals are effective control measures), demonstrates the value of continued surveillance of avian influenza viruses, and emphasizes the need for evaluation and integration of improved risk assessment measures.

Figure 4. HA Structural Analysis and Glycan Microarray Analysis

(A) Localization of amino acid changes identified in viruses recovered from ferrets in the transmission study. Shown is the 3D structure of the monomer of A/Brevig Mission/1/18 (H1N1) HA in complex with human receptor analogs (Protein Data Bank [PDB] ID code 2WRG). A close-up view of the globular head is also shown to the right. Mutations known to increase affinity to human-type receptors are shown in red (E190D and G225D). Mutations that emerged in HA during replication and/or transmission in ferrets are shown in green (E89D, S113N, I187T, and D265V). The amino acid changes at positions 89 and 113 are located close to an amino acid at position 110 (103 with H5 numbering) that was previously found to be associated with the transmissibility of an H5 virus (Herfst et al., 2012). Images were created with MacPyMOL (http://www.pymol.org/).

(B–E) The receptor specificities of viruses possessing 1918 HA (B), 1918-like avian HA (C), 1918-like avian HA-190D/225D (D), and 1918-like avian HA-89D/113N/190D/225D (E) were assessed by using a glycan microarray containing a diverse library of α2,3- and α2,6-linked sialosides (Xu et al., 2013). Viruses, directly labeled with biotin, were applied at 128 hemagglutination units/ml for 1 hr and, after washing, were incubated with Streptavidin-Alexa Fluor 647 (1 μg/ml) for 1 hr to detect bound virus. Error bars represent the SD calculated from six replicate spots of each glycan. A complete list of glycans is found in Table S6. See also Figure S6.



(legend on next page)

EXPERIMENTAL PROCEDURES

Cells and Viruses

Human embryonic kidney 293T cells and MDCK cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum (FCS) and in minimal essential medium (MEM) containing 5% newborn calf serum, respectively. HeLa cells were maintained in MEM containing 10% FCS. All cells were maintained at 37°C in 5% CO_2.

All viruses used in this study, except for A/duck/Alberta/35/76 (H1N1; DK/ALB), were generated by using plasmid-driven reverse genetics as described in the Supplemental Experimental Procedures and a previous report (Neumann et al., 1999). Virus growth in cells and eggs was examined as described in the Supplemental Experimental Procedures.

Experimental Infection of Mice and Ferrets

Female BALB/c mice (5–6 weeks old) (Jackson laboratory) were used to determine MLD_{50} values. Under anesthesia, four mice per group were intranasally inoculated with viruses. Body weight and survival were monitored daily for 14 days.

Female ferrets (5–8 months old) (Triple F Farms), which were serologically negative by hemagglutination inhibition (HI) assay for currently circulating human influenza viruses, were used in this study. Under anesthesia, six ferrets per group were intranasally inoculated with 10^6 pfu (0.5 ml) of viruses. Three ferrets per group were euthanized on days 3 and 6 postinfection for virological and pathological examinations. The virus titers in various organs were determined by use of plaque assays in MDCK cells. For the transmission study, pairs of ferrets were individually housed in adjacent wireframe cages (Showa Science) that prevented direct and indirect contact between the animals but allowed spread of influenza virus by respiratory droplets, and virus titers in nasal washes collected from each animal were determined by use of plaque assays in MDCK cells as described elsewhere (Imai et al., 2012) and in the Supplemental Experimental Procedures.

All experiments with mice and ferrets were performed in accordance with the University of Wisconsin-Madison's Regulations for Animal Care and Use and approved by the Animal Experiment Committee of the University of Wisconsin-Madison.

Glycan Arrays

Glycan array analysis was performed on a glass slide microarray containing 6 replicates of 58 diverse sialic acid-containing glycans including terminal sequences and intact *N*-linked and *O*-linked glycans found on mammalian and avian glycoproteins and glycolipids (Xu et al., 2013). Virus samples were directly labeled with biotin (Ramya et al., 2013) and then applied to the slide array; slide scanning to detect virus bound to glycans was conducted as described previously (Watanabe et al., 2013) and in the Supplemental Experimental Procedures. A complete list of glycans present on the array is provided in Table S6.

Pathological Examination

Excised tissues of animal organs preserved in 10% phosphate-buffered formalin were processed for paraffin embedding and cut into 3 μm thick sections. One section from each tissue sample was stained using a standard hematoxylin and eosin (H&E) procedure, whereas another one was processed for immunohistological staining with a rabbit polyclonal antibody for type A influenza nucleoprotein (NP) antigen (prepared in our laboratory) that reacts comparably with all of the viruses tested in this study. Specific antigen-antibody reactions were visualized with 3,3′-diaminobenzidine tetrahydrochloride staining by using the DAKO LSAB2 system (DAKOCytomation). Pathological scores were determined as described in the Supplemental Experimental Procedures.

Serological Tests

Human serum samples obtained from individuals vaccinated with the A/California/07/2009 (H1N1) hemagglutinin (HA) split vaccine in Japan were used for serological testing. Hemagglutination inhibition titers were determined as described in the Supplemental Experimental Procedures. All experiments with human sera were approved by the Research Ethics Review Committee of the Institute of Medical Science, the University of Tokyo (approval number 21-38-1117), and the Institutional Review Board of the University of Wisconsin-Madison.

Neuraminidase Inhibition Assay

The sensitivity of viral NA to oseltamivir carboxylate was evaluated by using an NA enzyme inhibition assay based on the methylumbelliferyl-N-acetylneuraminic acid (MUNANA) method as described in previous studies (Gubareva et al., 2001; Kiso et al., 2004) and in the Supplemental Experimental Procedures.

Polykaryon Formation Assay

HeLa cells were transfected with expression plasmids expressing various HAs. Polykaryon formation was induced by exposing cells to low pH buffer and was observed as described in the Supplemental Experimental Procedures.

Thermostability of HA

Viruses (128 hemagglutination units in PBS) were incubated at 50°C for the times indicated. Infectivity and hemagglutination activity were determined by using plaque assays in MDCK cells and the hemagglutination assay with 0.5% turkey blood cells, respectively.

Luciferase-Based Minigenome Assay

A luciferase-based minigenome assay was performed to examine viral polymerase activity as described previously (Ozawa et al., 2007). Briefly, 293T cells were transfected with plasmids for the expression of viral proteins PA, PB1, PB2, and NP derived from 1918-like avian virus or its mutants and with pPolI-WSN-NA-firefly-luciferase. Plasmid pGL4.74[hRuc/TK] (Promega) served as an internal control for the dual-luciferase assay. After transfection, the cells were incubated at 33°C or 37°C for 24 hr, and then luciferase activity was measured with the dual-luciferase reporter system (Promega) on a GloMax Microplate Luminometer (Promega) according to the manufacturer's instructions.

Statistical Analysis

Statistical analyses were performed by using ANOVA in GraphPad Prism version 5.0 (GraphPad); p values of < 0.05 were considered significant.

Biosafety and Biosecurity

All experiments with 1918-related viruses were performed in biosafety level 3 (BSL3) agriculture containment laboratories. In vitro experiments were conducted in Class II biological safety cabinets, and transmission experiments were conducted in HEPA-filtered ferret isolators (Imai et al., 2012). The research program, procedures, occupational health plan, documentation, security, and facilities are reviewed annually by the University of Wisconsin-Madison

Figure 5. Global Patterns of PB2 Proteins Derived from Avian Influenza Viruses

(A) Phylogenetic tree of 1,022 randomly selected amino acid sequences of PB2 genes derived from avian influenza viruses. The tree was rooted to the PB2 sequence of the 1918 virus, A/Brevig Mission/1/18 (H1N1) (red arrow). The PB2 from A/blue-winged teal/Ohio/926/2002 (H3N8), which is expressed by the 1918-like avian virus, is indicated by the blue arrow. The year in which the strains were isolated is indicated by horizontal bars to the right of the tree drawn at the same vertical position as the position of the strain in the tree. The tree and time series are color coded according to the number of amino acid differences from the 1918 virus defined in the histogram shown in (C).

(B) Geographic map indicating locations where the respective viruses shown in (A) were isolated. The map is color coded according to the number of amino acid differences defined in the histogram shown in (C).

(C) Histogram showing the distribution of the amino acid differences of all avian influenza PB2 proteins from the 1918 PB2 protein and color scheme for (A) and (B) (*red = 0–5 amino acid substitutions; magenta = 6–10; purple = 11–15*). See also Figure S6 and Tables S1 and S7.

Responsible Official and at regular intervals by the CDC and the Animal and Plant Health Inspection Service (APHIS) as part of the University of Wisconsin-Madison Select Agent Program. More detailed information on biosecurity and biosafety is described in the Supplemental Experimental Procedures.

Phylogenetic Analysis

For each viral gene segment, all nearly complete sequences from avian viruses were downloaded from the Influenza Research Database (http://www.fludb.org). These sequences were compared with the corresponding gene sequences of the 1918 virus to determine the number of amino acid sequence differences from the 1918 virus. Due to the large number of avian influenza gene sequences (n = 3,031–5,836 sequences for the PB2, PB1, PA, NP, M, and NS segments), roughly 1,000 sequences per gene were randomly selected to enable efficient phylogenetic analysis.

Phylogenetic trees for each segment were constructed with PhyML software package version 3.0 using subtree pruning and regrafting to optimize tree topology. The randomly selected subsets of amino acid sequences for each gene segment were aligned by using MAFFT (http://www.ebi.ac.uk/Tools/msa/mafft/).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.chom.2014.05.006.

AUTHOR CONTRIBUTIONS

T.W., G.N., and Y.K. designed the study; T.W., G.Z., M.H., A.H., R.M., S.F., Y.T., and S.W. performed the experiments; C.A.R., D.F.B., and D.J.S. conducted phylogenetic analysis; N.N., K.T., and H.H. conducted pathologic examination; T.W., G.Z., C.A.R., N.N., M.H., R.M., S.W., M.I., G.N., H.H., J.C.P., D.J.S., and Y.K. analyzed the data; T.W., C.A.R., N.N., R.M., E.A.M., S.W., M.I., G.N., H.H., J.C.P., and Y.K. wrote the manuscript; and Y.K. oversaw the project. T.W., G.Z., C.A.R. contributed equally to this work.

ACKNOWLEDGMENTS

We thank Kelly Moore, Naomi Fujimoto, Izumi Ishikawa, and Yuko Sato for technical support. We thank Dr. Susan Watson for editing the manuscript. Several glycans in the glycan array were provided by the Consortium for Functional Glycomics (http://www.functionalglycomics.org/) funded by NIGMS grant GM62116. This work was also supported by National Institute of Allergy and Infectious Diseases Public Health Service research grants, by RO1 AI080598 and R56 AI099275, by ERATO (Japan Science and Technology Agency), and by the Strategic Basic Research Programs of Japan Science and Technology Agency, Japan. C.A.R. was supported by a University Research Fellowship from the Royal Society.

Received: January 30, 2014
Revised: March 25, 2014
Accepted: April 24, 2014
Published: June 11, 2014

REFERENCES

Beare, A.S., and Webster, R.G. (1991). Replication of avian influenza viruses in humans. Arch. Virol. *119*, 37–42.

Belser, J.A., Gustin, K.M., Pearce, M.B., Maines, T.R., Zeng, H., Pappas, C., Sun, X., Carney, P.J., Villanueva, J.M., Stevens, J., et al. (2013). Pathogenesis and transmission of avian influenza A (H7N9) virus in ferrets and mice. Nature *501*, 556–559.

Chutinimitkul, S., Herfst, S., Steel, J., Lowen, A.C., Ye, J., van Riel, D., Schrauwen, E.J., Bestebroer, T.M., Koel, B., Burke, D.F., et al. (2010). Virulence-associated substitution D222G in the hemagglutinin of 2009 pandemic influenza A(H1N1) virus affects receptor binding. J. Virol. *84*, 11802–11813.

de Vries, R.P., Zhu, X., McBride, R., Rigter, A., Hanson, A., Zhong, G., Hatta, M., Xu, R., Yu, W., Kawaoka, Y., et al. (2014). Hemagglutinin receptor specificity and structural analyses of respiratory droplet-transmissible H5N1 viruses. J. Virol. *88*, 768–773.

Gabriel, G., Herwig, A., and Klenk, H.D. (2008). Interaction of polymerase subunit PB2 and NP with importin alpha1 is a determinant of host range of influenza A virus. PLoS Pathog. *4*, e11.

Gubareva, L.V., Kaiser, L., Matrosovich, M.N., Soo-Hoo, Y., and Hayden, F.G. (2001). Selection of influenza virus mutants in experimentally infected volunteers treated with oseltamivir. J. Infect. Dis. *183*, 523–531.

Hatta, M., Gao, P., Halfmann, P., and Kawaoka, Y. (2001). Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. Science *293*, 1840–1842.

Hatta, M., Hatta, Y., Kim, J.H., Watanabe, S., Shinya, K., Nguyen, T., Lien, P.S., Le, Q.M., and Kawaoka, Y. (2007). Growth of H5N1 influenza A viruses in the upper respiratory tracts of mice. PLoS Pathog. *3*, 1374–1379.

Herfst, S., Schrauwen, E.J., Linster, M., Chutinimitkul, S., de Wit, E., Munster, V.J., Sorrell, E.M., Bestebroer, T.M., Burke, D.F., Smith, D.J., et al. (2012). Airborne transmission of influenza A/H5N1 virus between ferrets. Science *336*, 1534–1541.

Hinshaw, V.S., Bean, W.J., Webster, R.G., Rehg, J.E., Fiorelli, P., Early, G., Geraci, J.R., and St Aubin, D.J. (1984). Are seals frequently infected with avian influenza viruses? J. Virol. *51*, 863–865.

Imai, M., Watanabe, T., Hatta, M., Das, S.C., Ozawa, M., Shinya, K., Zhong, G., Hanson, A., Katsura, H., Watanabe, S., et al. (2012). Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. Nature *486*, 420–428.

Johnson, N.P., and Mueller, J. (2002). Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. Bull. Hist. Med. *76*, 105–115.

Jones, K.E., Patel, N.G., Levy, M.A., Storeygard, A., Balk, D., Gittleman, J.L., and Daszak, P. (2008). Global trends in emerging infectious diseases. Nature *451*, 990–993.

Kiso, M., Mitamura, K., Sakai-Tagawa, Y., Shiraishi, K., Kawakami, C., Kimura, K., Hayden, F.G., Sugaya, N., and Kawaoka, Y. (2004). Resistant influenza A viruses in children treated with oseltamivir: descriptive study. Lancet *364*, 759–765.

Koçer, Z.A., Krauss, S., Stallknecht, D.E., Rehg, J.E., and Webster, R.G. (2012). The potential of avian H1N1 influenza A viruses to replicate and cause disease in mammalian models. PLoS ONE *7*, e41609.

Li, Z., Chen, H., Jiao, P., Deng, G., Tian, G., Li, Y., Hoffmann, E., Webster, R.G., Matsuoka, Y., and Yu, K. (2005). Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. J. Virol. *79*, 12058–12064.

Liu, Y., Childs, R.A., Matrosovich, T., Wharton, S., Palma, A.S., Chai, W., Daniels, R., Gregory, V., Uhlendorff, J., Kiso, M., et al. (2010). Altered receptor specificity and cell tropism of D222G hemagglutinin mutants isolated from fatal cases of pandemic A(H1N1) 2009 influenza virus. J. Virol. *84*, 12069–12074.

Matrosovich, M., Tuzikov, A., Bovin, N., Gambaryan, A., Klimov, A., Castrucci, M.R., Donatelli, I., and Kawaoka, Y. (2000). Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. J. Virol. *74*, 8502–8512.

Matrosovich, M.N., Gambaryan, A.S., and Klenk, H.-D. (2008). Receptor Specificity of Influenza Viruses and Its Alteration during Interspecies Transmission. In Avian Influenza, H.-D. Klenk, M.N. Matrosovich, and J. Stech, eds. (Basel: Karger), pp. 134–155.

Murphy, B.R., Buckler-White, A.J., London, W.T., Harper, J., Tierney, E.L., Miller, N.T., Reck, L.J., Chanock, R.M., and Hinshaw, V.S. (1984). Avian-human reassortant influenza A viruses derived by mating avian and human influenza A viruses. J. Infect. Dis. *150*, 841–850.

Neumann, G., and Kawaoka, Y. (2006). Host range restriction and pathogenicity in the context of influenza pandemic. Emerg. Infect. Dis. *12*, 881–886.

Neumann, G., Watanabe, T., Ito, H., Watanabe, S., Goto, H., Gao, P., Hughes, M., Perez, D.R., Donis, R., Hoffmann, E., et al. (1999). Generation of influenza

A viruses entirely from cloned cDNAs. Proc. Natl. Acad. Sci. USA *96*, 9345–9350.

Ozawa, M., Fujii, K., Muramoto, Y., Yamada, S., Yamayoshi, S., Takada, A., Goto, H., Horimoto, T., and Kawaoka, Y. (2007). Contributions of two nuclear localization signals of influenza A virus nucleoprotein to viral replication. J. Virol. *81*, 30–41.

Rabadan, R., Levine, A.J., and Robins, H. (2006). Comparison of avian and human influenza A viruses reveals a mutational bias on the viral genomes. J. Virol. *80*, 11887–11891.

Ramya, T.N., Weerapana, E., Cravatt, B.F., and Paulson, J.C. (2013). Glycoproteomics enabled by tagging sialic acid- or galactose-terminated glycans. *Glycobiology 23*, 211–221.

Reid, A.H., Taubenberger, J.K., and Fanning, T.G. (2004). Evidence of an absence: the genetic origins of the 1918 pandemic influenza virus. Nat. Rev. Microbiol. *2*, 909–914.

Richard, M., Schrauwen, E.J., de Graaf, M., Bestebroer, T.M., Spronken, M.I., van Boheemen, S., de Meulder, D., Lexmond, P., Linster, M., Herfst, S., et al. (2013). Limited airborne transmission of H7N9 influenza A virus between ferrets. Nature *501*, 560–563.

Schnirring, L. (2014). Eleven new H7N9 cases include one in Beijing. Center for Infectious Disease Research and Policy, University of Minnesota. http://www.cidrap.umn.edu/news-perspective/2014/01/eleven-new-h7n9-cases-include-one-beijing.

Smith, G.J., Bahl, J., Vijaykrishna, D., Zhang, J., Poon, L.L., Chen, H., Webster, R.G., Peiris, J.S., and Guan, Y. (2009). Dating the emergence of pandemic influenza viruses. Proc. Natl. Acad. Sci. USA *106*, 11709–11712.

Subbarao, E.K., London, W., and Murphy, B.R. (1993). A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. J. Virol. *67*, 1761–1764.

Taubenberger, J.K., and Morens, D.M. (2006). 1918 Influenza: the mother of all pandemics. Emerg. Infect. Dis. *12*, 15–22.

Taylor, L.H., Latham, S.M., and Woolhouse, M.E. (2001). Risk factors for human disease emergence. Philos. Trans. R. Soc. Lond. B Biol. Sci. *356*, 983–989.

Tumpey, T.M., García-Sastre, A., Mikulasova, A., Taubenberger, J.K., Swayne, D.E., Palese, P., and Basler, C.F. (2002). Existing antivirals are effective against influenza viruses with genes from the 1918 pandemic virus. Proc. Natl. Acad. Sci. USA *99*, 13849–13854.

Van Hoeven, N., Pappas, C., Belser, J.A., Maines, T.R., Zeng, H., García-Sastre, A., Sasisekharan, R., Katz, J.M., and Tumpey, T.M. (2009). Human HA and polymerase subunit PB2 proteins confer transmission of an avian influenza virus through the air. Proc. Natl. Acad. Sci. USA *106*, 3366–3371.

Watanabe, T., Shinya, K., Watanabe, S., Imai, M., Hatta, M., Li, C., Wolter, B.F., Neumann, G., Hanson, A., Ozawa, M., et al. (2011). Avian-type receptor-binding ability can increase influenza virus pathogenicity in macaques. J. Virol. *85*, 13195–13203.

Watanabe, T., Kiso, M., Fukuyama, S., Nakajima, N., Imai, M., Yamada, S., Murakami, S., Yamayoshi, S., Iwatsuki-Horimoto, K., Sakoda, Y., et al. (2013). Characterization of H7N9 influenza A viruses isolated from humans. Nature *501*, 551–555.

Webster, R.G., Yakhno, M., Hinshaw, V.S., Bean, W.J., and Murti, K.G. (1978). Intestinal influenza: replication and characterization of influenza viruses in ducks. Virology *84*, 268–278.

Xu, R., Krause, J.C., McBride, R., Paulson, J.C., Crowe, J.E., Jr., and Wilson, I.A. (2013). A recurring motif for antibody recognition of the receptor-binding site of influenza hemagglutinin. Nat. Struct. Mol. Biol. *20*, 363–370.

Yang, H., Carney, P., and Stevens, J. (2010). Structure and Receptor binding properties of a pandemic H1N1 virus hemagglutinin. PLoS Curr. *2*, RRN1152.

Zhang, Q., Shi, J., Deng, G., Guo, J., Zeng, X., He, X., Kong, H., Gu, C., Li, X., Liu, J., et al. (2013). H7N9 influenza viruses are transmissible in ferrets by respiratory droplet. Science *341*, 410–414.

Zhu, H., Wang, D., Kelvin, D.J., Li, L., Zheng, Z., Yoon, S.W., Wong, S.S., Farooqui, A., Wang, J., Banner, D., et al. (2013). Infectivity, transmission, and pathology of human-isolated H7N9 influenza virus in ferrets and pigs. Science *341*, 183–186.