

ARTICLE

Received 6 Jan 2015 | Accepted 14 May 2015 | Published 17 Jun 2015

DOI: 10.1038/ncomms8491

OPEN

Identification of mammalian-adapting mutations in the polymerase complex of an avian H5N1 influenza virus

Andrew S. Taft¹, Makoto Ozawa^{2,3}, Adam Fitch⁴, Jay V. Depasse⁴, Peter J. Halfmann¹, Lindsay Hill-Batorski¹, Masato Hatta¹, Thomas C. Friedrich^{5,6}, Tiago J.S. Lopes^{1,9}, Eileen A. Maher¹, Elodie Ghedin^{4,7}, Catherine A. Macken⁸, Gabriele Neumann¹ & Yoshihiro Kawaoka^{1,9,10}

Avian influenza viruses of the H5N1 subtype pose a serious global health threat due to the high mortality (>60%) associated with the disease caused by these viruses and the lack of protective antibodies to these viruses in the general population. The factors that enable avian H5N1 influenza viruses to replicate in humans are not completely understood. Here we use a high-throughput screening approach to identify novel mutations in the polymerase genes of an avian H5N1 virus that confer efficient polymerase activity in mammalian cells. Several of the identified mutations (which have previously been found in natural isolates) increase viral replication in mammalian cells and virulence in infected mice compared with the wild-type virus. The identification of amino-acid mutations in avian H5N1 influenza virus polymerase complexes that confer increased replication and virulence in mammals is important for the identification of circulating H5N1 viruses with an increased potential to infect humans.

¹Influenza Research Institute, School of Veterinary Medicine, Department of Pathobiological Sciences, University of Wisconsin-Madison, Madison, Wisconsin 53711, USA. ²Laboratory of Animal Hygiene, Joint Faculty of Veterinary Medicine, Kagoshima University, Kagoshima 890-0065, Japan. ³Transboundary Animal Diseases Center, Joint Faculty of Veterinary Medicine, Kagoshima University, Kagoshima 890-0065, Japan. ⁴University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261, USA. ⁵Wisconsin National Primate Research Center, Madison Wisconsin 53715, USA. ⁶Department of Pathobiological Sciences, University of Wisconsin-Madison, Madison WI 53711, USA. ⁷Department of Biology, New York University, New York, New York 10003, USA. ⁸Bioinformatics Institute, University of Auckland, Auckland 1010, New Zealand. ⁹Division of Virology, Department of Microbiology and Immunology and International Research Center for Infectious Diseases, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan. ¹⁰Infection-Induced Host Responses Project, Exploratory Research for Advanced Technology, Saitama 332-0012, Japan. Correspondence and requests for materials should be addressed to Y.K. (email: kawaokay@svm.vetmed.wisc.edu)

Influenza A viruses contain eight negative-sense viral RNA (vRNA) segments that upon being transported to the nucleus serve as templates for gene replication and the synthesis of viral mRNAs. The replication of vRNAs and the transcription of viral mRNAs are catalysed by the viral polymerase complex, which is composed of the PB2, PB1 and PA subunits. The viral RNA-dependent RNA polymerase activity is encoded by the PB1 protein. The PB2 protein binds to the cap structure of host pre-mRNAs; the endonuclease activity of PA then cleaves the cap structure and the adjacent nucleotides from cellular pre-mRNAs, thereby creating short, capped mRNA primers that are elongated by PB1 for viral mRNA synthesis. All three polymerase subunits, together with the nucleoprotein (NP), are required for efficient influenza virus replication.

Migratory waterfowl serves as the reservoir for influenza A viruses, which sporadically transmit to domestic poultry, pigs or other species; occasionally, influenza viruses originating from avian species ('avian influenza viruses') also infect humans. In recent years, there have been large-scale outbreaks of avian influenza viruses in humans. Since their emergence in China in 1996, highly pathogenic avian H5N1 viruses have caused 826 confirmed cases of human infection, resulting in 440 fatalities (http://www.who.int/influenza/human_animal_interface/H5N1_cumulative_table_archives/en/ data as of 31 March 2015). In 2013, H7N9 influenza viruses emerged in China and have since infected >571 individuals, causing 212 deaths (http://www.who.int/entity/influenza/human_animal_interface/influenza_h7n9/RiskAssessment_H7N9_23Feb20115.pdf; data as of 23 February 2015). Rare cases of human infection have also been caused by avian influenza viruses of the H9N2, H6N1, H7N7, H10N8, H7N2 and H7N3 subtypes^{1–6}; of these, human infections with H7N7 and H10N8 viruses have been fatal.

The haemagglutinin (HA) surface glycoprotein and the polymerase complex affect the host range of influenza viruses^{7–12}. The HA proteins of avian influenza viruses bind preferentially to α 2,3-linked sialic acids, which are predominantly expressed on epithelial cells in the intestinal tract of aquatic birds⁸. By contrast, human influenza viruses bind preferentially to α 2,6-linked sialic acids, which are prevalent on epithelial cells in the upper human airway⁸. In addition to HA-mediated host range restriction, the polymerase complex of avian influenza viruses is often restricted in its replicative ability in mammalian cells. This host range restriction may be a consequence of polymerase interactions with host-specific cellular proteins, and/or temperature differences between the intestinal tract of birds (38–41 °C) and the upper respiratory tract of mammals (33 °C). We¹³ and others¹⁴ identified the amino acid at position 627 of PB2 as a major host range determinant of influenza A viruses. In addition, the amino acids at positions 591 (refs 15,16) and 701 (refs 17,18) of PB2 affect the replication of avian influenza viruses in mammals. Moreover, other mutations in PB2, PB1 and PA also influence the host range of influenza viruses; examples include PB2-147T/339T/588T (ref. 19), PB2-271A (ref. 20), PB2-158G (ref. 21), PB1-13L/678S (ref. 18) and PA-97I (ref. 22).

The significance of the influenza viral polymerase complex for host range restriction is now well established, primarily based on studies that compared viruses of high or low virulence in mammals. Although such studies led to the identification of host range determinants such as PB2-627, -591 and -701N, there may be additional, currently unknown amino-acid changes in the polymerase proteins that affect host range. The testing of all possible amino-acid changes in the polymerase proteins is not feasible; hence, no systematic efforts have been undertaken to catalogue potentially mammalian-adapting mutations in the viral polymerase complex. Therefore, in the present study, we implemented a high-throughput, reporter gene-based screening

approach that allows the testing of large numbers of polymerase mutants. By using this approach, we identified several mutations in an avian H5N1 polymerase complex that increase polymerase activity in mammalian cells and virulence in mice.

Results

Establishment of a high-throughput screening approach. First, we established a high-throughput screening system to test large numbers of viral polymerase mutants for their ability to confer efficient replication to avian H5N1 viral polymerase complexes in mammalian cells. We generated a biologically contained virus that possesses the viral polymerase (PB2, PB1 and PA) and NP RNA segments of the avian H5N1 A/Muscovy duck/Vietnam/TY93/2007 (TY93/H5N1) virus (Clade 2.3.4), the viral NA, M and NS RNA segments of the laboratory-adapted A/WSN/33 (WSN; H1N1) virus, and a modified WSN HA RNA segment in which most of the open reading frame of HA was replaced with that of the reporter protein green fluorescent protein (GFP; essentially as described before in ref. 23; Fig. 1a). The replication of this virus is restricted to a cell line that stably expresses the HA protein. The TY93/H5N1 PB2 protein of this virus (termed TY93/H5N1 GFP-627E) encodes glutamic acid at position 627, which restricts the replicative ability of avian H5N1 virus polymerase complexes in mammals¹³. As a control virus, we generated TY93/H5N1 GFP-627K encoding the PB2-E627K mutation known to increase the replicative ability of avian influenza virus polymerase complexes in mammalian cells¹³. The replication-incompetent TY93/H5N1 GFP-627E and -E627K viruses can be propagated in cell lines expressing the WSN-HA protein, but do not replicate in wild-type cells (Fig. 1b). The University of Wisconsin-Madison Institutional Biosafety Committee (IBC) approved work with these viruses in BSL-2 containment.

To test the replication-incompetent, GFP-expressing viruses for high-throughput screening approaches in human cells, we inoculated human embryonic fibroblast (293) cells stably expressing the WSN-HA protein (293-HA) with both viruses at a multiplicity of infection (MOI) of 0.1. Five hours later, we subjected cells to fluorescence-activated cell sorting (FACS) analysis. As expected, TY93/H5N1 GFP-627E replication was restricted in human 293-HA cells, resulting in low levels of GFP expression; by contrast, the PB2-E627K mutation conferred efficient replication and higher levels of GFP expression to TY93/H5N1 GFP in human 293-HA cells (Supplementary Fig. 1a,b). Hence, this GFP-based approach can be used to identify polymerase mutants with increased replicative ability in human cells.

Generation of mutant virus libraries. To identify mutations that increase the replicative ability of the TY93/H5N1 polymerase complex in mammalian cells, we generated 'virus libraries' possessing random mutations in the TY93/H5N1 polymerase proteins. To generate large numbers of mutants, each polymerase gene was divided into four regions of ~150–200 amino acids; for example, we divided the PA gene into regions encoding amino acids 1–186, 187–328, 329–574 and 575–716 (Fig. 1c). All 12 polymerase fragments were amplified with an error-prone polymerase, resulting in 1–2 random amino-acid changes per amplified region. For the PB1 and PA genes, we then mixed the four randomly mutated complementary DNA libraries of each gene, resulting in PA and PB1 plasmid libraries. For PB2, we mixed the three libraries possessing mutations at positions 1–207, 208–394 and 395–587, resulting in a PB2 plasmid library with mutations at amino-acid positions 1–587. The C-terminal library encoding mutations at positions 588–759 of PB2 was not combined with the other three PB2 libraries

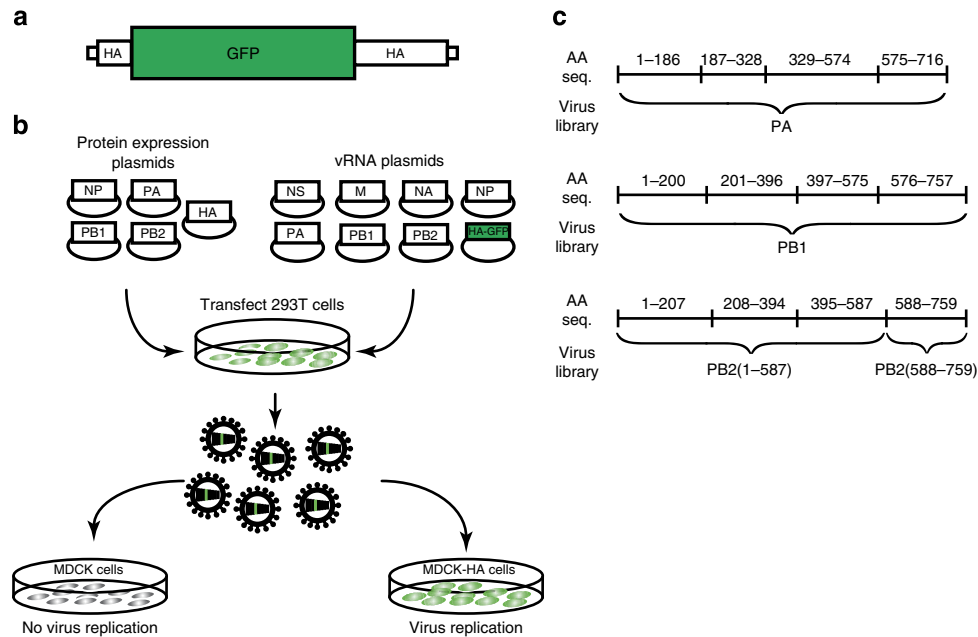


Figure 1 | Schematic overview of HA gene-deficient influenza virus used for high-throughput screens. (a) Schematic diagram of influenza virus-like RNA encoding the reporter protein GFP. The GFP coding region is flanked by viral HA sequences that are required for efficient virion incorporation of the vRNA, and by viral regulatory sequences at both ends (indicated by small bars). **(b)** Generation of HA gene-deficient influenza virus. 293T cells were transfected with plasmids expressing HA, the viral polymerase (PB2, PB1 and PA) proteins and the NP. Cells were co-transfected with plasmids synthesizing all eight vRNAs; the plasmid for the synthesis of wild-type HA was replaced with that encoding GFP **(a)**. Influenza virus encoding GFP can be propagated in MDCK cells expressing HA, but not in normal MDCK cells. **(c)** Summary of plasmid and virus libraries. AA seq., amino-acid sequence.

because many of the mutations known to enhance the polymerase activity in mammalian cells (such as PB2-Q591K, -E627K and D701N) are located in this region. By using reverse genetics (Fig. 1b), we then used the complementary DNA libraries to generate four different virus libraries that possess several random amino-acid mutations in PB2(1–587), PB2(588–759), PB1 and PA (Fig. 2). All mutant virus libraries were amplified in 293-HA cells.

Pilot screen to identify mutations that confer increased polymerase activity. To test our experimental approach, we first performed a pilot screen with the TY93/H5N1 PB2(588–759) virus library, which was chosen because it encompasses PB2 amino-acid positions that are known to be essential for the mammalian adaptation of avian influenza viruses^{13,15,18}. We infected 293-HA cells with the TY93/H5N1 PB2(588–759) virus library at an MOI of 0.1 and performed FACS analysis from 4–6 h post infection (Supplementary Fig. 1d). Since viral protein synthesis (and therefore GFP expression) increased during this time period, we recalibrated the sorting gate at 5 and 6 h post infection. (Re)Calibrations were carried out by analysing 10,000 cells infected with TY93/H5N1 GFP-627E virus (Supplementary Fig. 1c). We analysed $> 8 \times 10^6$ cells, and isolated 3,744 cells with GFP expression levels higher than those detected for TY93/H5N1 GFP-627E virus (Fig. 2b). These virus-infected cells were then sorted individually into wells of a 96-well plate containing Madin–Darby canine kidney (MDCK) cells stably expressing WSN-HA (MDCK-HA). Seventy-two hours later, we measured the GFP expression levels of the infected cells and detected 1,485 GFP-positive samples. For these positive samples, the virus-containing supernatants were harvested and stored for further analysis.

Next, we sequenced (by using the Sanger sequencing method) the PB2 genes of the 96 virus samples with the highest GFP

expression levels. All but 11 samples possessed mutations in PB2, including the known mammalian-adapting PB2-E627K change (Supplementary Table 1), which was found in 10 samples. Interestingly, the most frequently detected mutation was PB2-627V, which has not previously been reported to be associated with the adaptation of avian influenza viruses to mammals. The identification of known and potentially novel markers of avian influenza virus adaptation to humans established the feasibility of our experimental platform.

Screens to identify mutations that confer increased polymerase activity. We next screened the remaining three mutant virus libraries and re-screened the TY93/H5N1 PB2(588–759) mutant virus library. In total, we screened $> 8 \times 10^6$ cells and detected 3,168 cells with increased GFP expression levels relative to those of TY93/H5N1 GFP-627E virus (Fig. 2b). Almost half of these cells (that is, 1,486 cells) resulted from infection with the TY93/H5N1 PB2(588–759) library, with proportionally fewer candidates being isolated after infection with the other three libraries. This finding may indicate that functions in the polymerase complex associated with mammalian adaptation are primarily encoded in the C-terminal portion of PB2, consistent with earlier findings^{13–15,24}. After individually incubating the 3,168 virus-infected, GFP-positive cells in MDCK-HA layers for 48 h, 713 samples were found to be GFP positive; the remaining cell samples may have died during the FACS screen, given that there was no visible cytopathic effect in the cell monolayers (Fig. 2b).

Because of the large number of potential mutant viruses recovered from our GFP-FACS screen, we conducted a second screen to assess the replicative ability of the viral polymerase complex by measuring the vRNA levels of the M gene. Briefly, we infected 293-HA cells with TY93/H5N1 GFP-627E, TY93/H5N1 GFP-627K and the 713 viruses isolated in our screens. Six hours later, we performed quantitative reverse transcription PCR

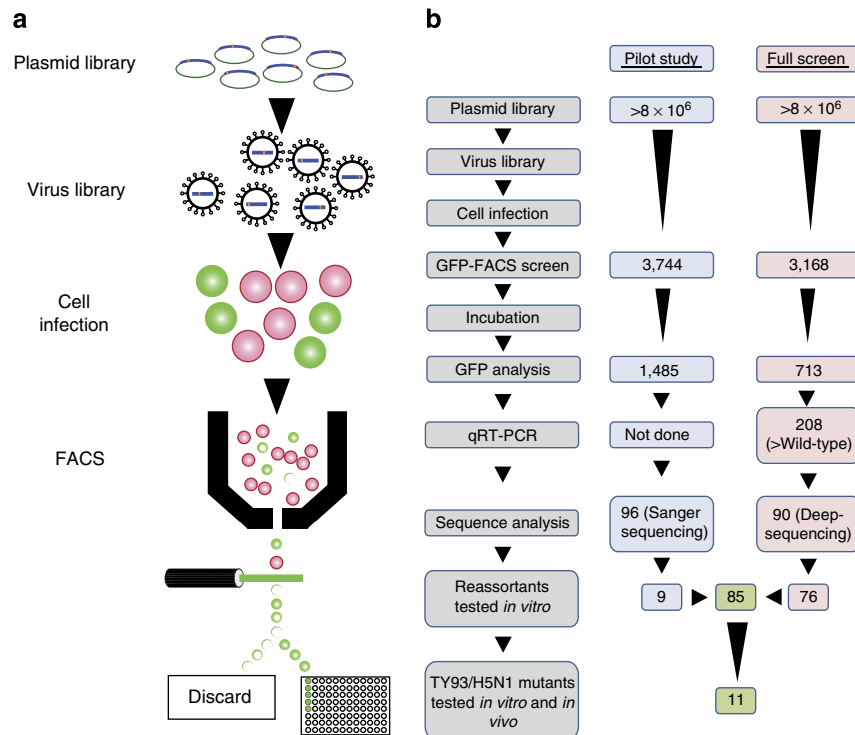


Figure 2 | Schematic overview of polymerase screens and characterization of selected mutations. (a) Schematic overview of polymerase screens. Plasmid libraries possessing random mutations in PB2(1–587), PB2(588–759), PB1 or PA were used to generate the respective HA-deficient virus libraries. 293 Cells expressing HA were infected with the virus libraries and screened by FACS analysis for cells with increased levels of GFP expression. Individual GFP-positive cells were incubated and retested for GFP expression. (b) Schematic overview of the number of viruses and mutants analysed. FACS screens yielded 3,744 and 3,168 GFP-positive samples for the pilot and full screens, respectively; virus amplification and GFP expression were confirmed for 1,485 and 713 of these. The latter samples were further analysed by qRT-PCR analysis: we detected 208 viruses with M vRNA levels higher than those of the control virus. A total of 90 candidates with the highest increases in viral replication were characterized by deep sequencing of the polymerase and NP genes. For the pilot study, 96 randomly selected samples were analysed by Sanger sequencing. From the mutants identified in the pilot and full screens, 85 were characterized in viral replication assays in cultured cells. Of those, 11 mutations were introduced into authentic TY93/H5N1 virus and tested for their effects on replication in human and avian cells, and for virulence in mice.

(qRT-PCR) analysis with oligonucleotides specific to the viral M gene. During the first 6 h of infection, the M vRNA levels of TY93/H5N1 GFP-627E virus increased, on average, 2.5-fold; by contrast, we detected an ~40-fold average increase in M vRNA levels upon infection with TY93/H5N1 GFP-627K. For the GFP-positive viruses identified through GFP-FACS analysis, the increases in M vRNA levels ranged from no appreciable change to ~300-fold (Supplementary Data 1); for 208 samples, we detected M vRNA levels higher than those of TY93/H5N1 GFP-627E (Fig. 2b).

Deep-sequencing analysis of mutants with high GFP-expressing levels. Based on the level of increase in M RNA synthesis between 0 and 6 h post infection, we selected the top 20 candidates each from the TY93/H5N1 PA, PB1 and PB2(1–587) mutant virus library screens, and the top 30 candidates from the TY93/H5N1 PB2(588–759) mutant virus library screen for further analysis. For these 90 viruses, we performed deep-sequencing analysis of the polymerase and NP genes; NP was included because it also affects influenza virus replication²⁵. We focused on mutations that were found in at least 30% of the sequence reads, which were detected in 66 viruses, namely in 11/20, 30/30, 12/20 and 13/20 of the viruses isolated from the TY93/H5N1 PB2(1–587), PB2(588–759), PB1 and PA mutant virus libraries (Supplementary Tables 2–5). The remaining viruses (selected on the basis of M RNA levels higher than those detected for the parental virus) may possess mutations in the remaining viral genes not analysed here. Most of the mutations detected in

the TY93/H5N1 PB2(588–759) library affected the amino acid at position 627; the known mammalian-adapting mutation PB2-E627K was found in 13 of 30 viruses, and the PB2-E627V mutation (the most frequently detected mutation in the pilot study) was found in 9 of 30 viruses. In addition, we detected the known mammalian-adapting mutations PB2-Q591K and PB2-D701N in one and three virus samples, respectively. Interestingly, an appreciable number of viruses possessed mutations in regions not targeted by random mutagenesis (for example, viruses isolated from the TY93/H5N1 PB1 virus library also possessed mutations in NP or PA); in fact, several viruses isolated from the TY93/H5N1 PB1 or PA libraries did not possess mutations in these genes, but rather in NP, albeit at a relatively low frequency (Supplementary Tables 4 and 5). Mutations outside the region targeted for mutagenesis may have emerged during replication in mammalian cells, and may thus play a role in the adaptation of avian virus polymerase complexes to mammalian cells. In samples with >1 mutation, some mutations occurred with similar frequency, whereas the prevalence of others differed greatly. For example, sample no. 13 of the TY93/H5N1 PA library encodes PA T97I and Y232C mutations at a prevalence of 87% and 85%, respectively; in addition, this virus sample possesses PA F612I and S648N mutations that were found in 30% of the sequence reads (Supplementary Table 5). This finding may indicate that the T97I and Y232C mutations arose first, followed by the F612I and S648N mutations. These findings might also suggest a functional relationship between T97I and Y232C, and between F612I and S648N.

Replicative ability of selected mutants in the genetic background of a low-pathogenic H5N1 variant. Our screens identified an appreciable number of mutations in the polymerase and NP proteins of TY93/H5N1. For further analyses, we focused on mutations that fulfilled at least one of the following criteria: (i) Mutations with a prevalence of $\geq 30\%$ based on deep-sequencing analysis (Supplementary Tables 2–5); (ii) Mutations identified in human H5N1 virus sequences deposited in the Influenza Research Database (<http://www.fludb.org>; accessed 1 September 2012); or (iii) Mutations located outside the region targeted for mutagenesis (for example, a mutation in PA isolated from the TY93/H5N1 PB1 mutant virus library). Based on these criteria, we selected 96 single, double or triple mutations in the polymerase and NP genes (Supplementary Data 1) for further characterization. Our GFP-FACS screens were carried out with HA-deficient virus maintained in HA-expressing cells. To test the effects of the selected mutations on influenza viruses that could replicate in normal cells, we introduced the mutations into a recombinant H5N1 virus. For biosafety reasons, the recombinant H5N1 virus possessed the TY93/H5N1 polymerase and NP genes in combination with the HA, NA, M and NS genes from A/Vietnam/1203/2004 (H5N1) virus; the HA gene of this virus has been modified to replace the multibasic sequence at the HA cleavage site with a single basic amino acid that renders the virus low pathogenic. We were able to introduce 85 of the selected mutations into replicating virus; however, 11 mutants could not be generated (Supplementary Data 1). To test the replicative ability of the 85 mutant viruses in mammalian cells, we infected human 293 cells (used because the GFP-FACS screens were carried out in this cell line) at an MOI of 0.001, and incubated them at 33 and 37 °C (that is, the respective temperatures of the upper and lower respiratory tract of mammals). At 24, 48 and 72 h post infection, supernatants were harvested and viral replication was assessed by means of plaque assays in MDCK cells. Most mutations that conferred increased replication to the recombinant H5N1 virus in human cells were located in the C-terminal portion of PB2 (Supplementary Data 1), consistent with the role of this region in the host range and virulence of influenza viruses^{15,18,26}. Among these mutations were the known mammalian-adapting PB2-Q591K, -E627K and -D701N changes; however, we also identified additional mutations in this region (such as PB2-E627V and -D701V) that increased the replicative ability of the recombinant H5N1 virus in human cells. Moreover, several mutations in PB1 and PA increased the replication of the

H5N1 virus in human cells at 33 and/or 37 °C (Supplementary Data 1). Based on the viral growth properties in human 293 cells, the following 11 mutations were selected for further characterization: PA-T97I, PB1-N105S, PB2-E192K, -Q591K, -E627K, -E627V, -D701N, -D701V, -K702R, -D740N and -I758T (the PB2-Q591K, -E627K and -D701N mutations are known to enhance the replicative ability of avian H5N1 viruses in mammals^{13–17,27} but were included here for comparison with the potentially novel markers of mammalian adaptation).

Growth characteristics of mutant TY93/H5N1 viruses in mammalian and avian cells. To further assess the biological significance of the 11 selected mutations, we introduced them into authentic TY93/H5N1 virus and measured viral replication kinetics in three human cell lines: 293 cells, human lung carcinoma (A549) cells and human bronchial epithelial (Calu-3) cells (Fig. 3 and Supplementary Fig. 2). Wild-type TY93/H5N1 virus did not replicate to detectable levels in A549 cells at 33 °C (Fig. 3a), underscoring the limited ability of avian H5N1 viruses to replicate in human cells at the temperature of the upper respiratory tract. By contrast, several mutations conferred this ability, demonstrating their growth-enhancing effect in mammalian cells. At 37 °C, the PB2-627V and -701V mutants replicated to significantly higher titres than wild-type TY93/H5N1 virus at 24 h post infection (Fig. 3a and Table 1). The PB2-E627K and PB2-D701N mutations are known to facilitate adaptation of avian influenza viruses to mammals^{13,18,26}, but the replication-enhancing effect in mammalian cells of the PB2-627V and -701V residues had not been described. In 293 and Calu-3 cells, TY93/H5N1 virus replicated to appreciable levels at 33 and 37 °C (Supplementary Fig. 2); similar to results obtained in A549 cells, most mutations increased the replicative ability of TY93/H5N1 at one or both temperatures tested (Supplementary Fig. 2 and Table 1).

The effect of the mutations tested here could be global or limited to the human cell types used. To distinguish between these possibilities, we also assessed virus replication in chicken fibroblast DF-1 cells at 37 °C. In contrast to human cells, the replicative abilities of the mutant viruses were equal or attenuated compared with wild-type TY93/H5N1 virus (Fig. 3b and Table 1). Hence, the mutations identified here increased avian influenza virus replication in mammalian, but not avian cells, suggesting a role for these mutations in avian influenza virus adaptation to mammals.

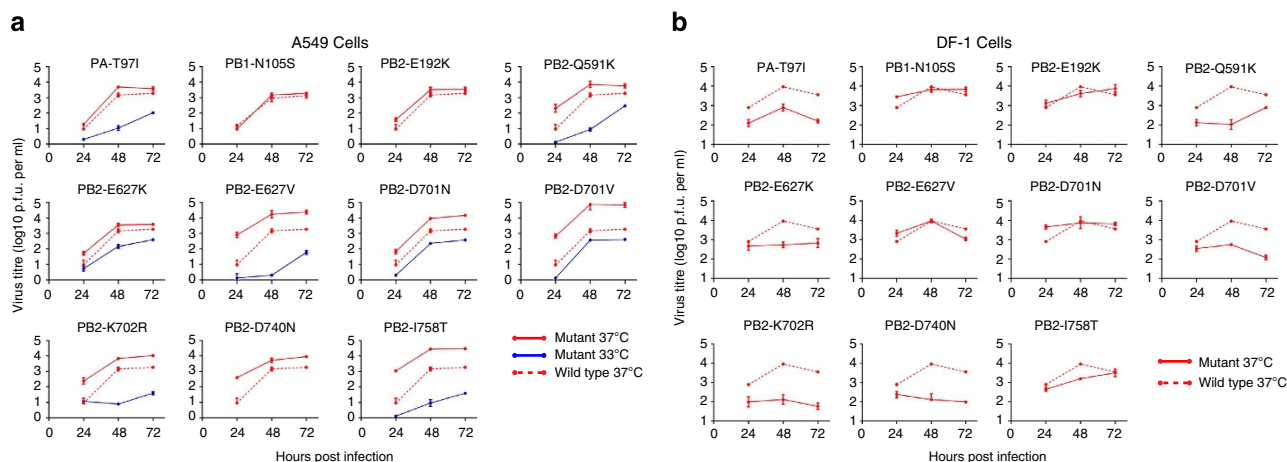


Figure 3 | Growth kinetics of mutant TY93/H5N1 viruses in mammalian and avian cells. Human A549 (a) or avian DF-1 (b) cells were infected with virus at a MOI of 0.001 and incubated at 33 or 37 °C. At the indicated time points post infection, virus titres were determined by use of plaque assays in MDCK cells. Values shown are the means (\pm s.d.) of three separate infections.

Table 1 | Replicative ability in human and avian cells and mouse virus titres of mutant TY93/H5N1 viruses.

Virus	In vitro replication*								In vitro polymerase activity†				Mouse virus titres*						MLD ₅₀ ‡
	293		A549		Calu-3		DF-1	293T		DF-1		3 Days post infection			5 Days post infection				
	33°C	37°C	33°C	37°C	33°C	37°C	37°C	33°C	37°C	33°C	37°C	Lung	Brain	N.T.	Lung	Brain	N.T.		
PA-I971	+	+	+	+	+	+	-	+++	+	=	=	=	=	=	=	=	=	17.8	
PB1-N105S	++	+	=	=	+	+	=	+	+	=	=	=	=	++	=	=	=	<1	
PB2-E192K	++	=	=	+	+	++	=	+++	+	++	+	=	++	++	=	=	=	<10	
PB2-Q591K	+	++	+	+	+	=	-	+++	+	+	+	=	=	++	=	=	=	18	
PB2-E627K	++	++	++	=	+	++	-	++++	+	+	+	=	+++	+++	=	=	=	18	
PB2-E627V	++	++	+	++	+	++	=	+++	+	=	=	=	=	++	=	=	=	18	
PB2-D701N	++	++	++	+	+	+	=	++	+	=	-	+	++	++	=	++	=	3.2	
PB2D-701V	+	+	++	++	++	++	-	++	=	=	-	+	++	++	=	++	=	3.2	
PB2-K702R	=	+	+	+	=	+	=	=	=	=	=	=	=	=	=	=	=	25	
PB2-D740N	+	+	=	+	=	+	-	=	=	+	=	=	=	=	=	=	=	436	
PB2-I758T	=	=	+	++	+	++	-	=	+	+	+	=	=	=	=	=	=	31	

N.T., nasal turbinates; p.f.u., plaque-forming units.

*Relative comparisons with wild-type virus titres: +, 2-10-fold increase; ++, >10-100-fold increase; +++ >100-fold increase; =, similar to wild-type virus (<2-fold change); -, titre lower than that of wild-type virus.

†Relative comparisons with wild-type polymerase activity: +, 2-25-fold increase; ++, >25-100-fold increase; +++, >100-500-fold increase; +++++, >500-fold increase; =, similar to wild-type (<2-fold change); -, activity lower than that of wild-type virus.

‡MLD₅₀ values are listed as p.f.u.; the MLD₅₀ value of the TY93/H5N1 wild-type virus is 178 p.f.u.

In vitro polymerase activity of polymerase mutants in mammalian and avian cells. To gain further insights into the effects of the selected mutations on the replicative ability of the mutant polymerase proteins, we utilized a minigenome assay in which cells are transfected with plasmids for the expression of wild-type or mutant viral polymerase and NP proteins, and with a plasmid for the synthesis of a virus-like RNA encoding luciferase (Fig. 4). In mammalian cells, the higher virus titres detected for these mutants (Fig. 3 and Supplementary Fig. 2) typically correlated with higher polymerase activity in minireplicon assays at 33 and/or 37 °C (Fig. 4a,b and Table 1). The PB2-K702R and -D740N mutations did not have appreciable effects on polymerase activity in minireplicon assays, even though they conferred higher virus titres in A549 cells (Fig. 3 and Table 1). The relative increases in mammalian cells compared with wild-type TY93/H5N1 replicative ability were greater at 33 °C than at 37 °C, most likely as a consequence of the low polymerase activity of TY93/H5N1 in mammalian cells at 33 °C (Fig. 3 and Supplementary Fig. 2). In avian DF-1 cells, several mutations including PB2-E192K, -Q591K, -E627K and -D740N conferred increased polymerase activity at 33 °C compared with TY93/H5N1 (Fig. 4 and Table 1); however, these increases in polymerase activity did not result in increased virus titres in DF-1 cells (Fig. 3b and Table 1). Together, these data indicate that there is not always a strict correlation between replicative ability in minireplicon assays (which assess viral RNA replication and transcription) and virus titres (which reflect the entire viral life cycle). This is consistent with our earlier findings¹⁹ and suggests that some of the mutations may affect steps in the viral life cycle other than vRNA transcription and/or replication, such as the nuclear export of viral ribonucleoprotein complexes.

Viral replication and pathogenicity in mice. Our *in vitro* data identified several mutations in the TY93/H5N1 polymerase proteins that increased virus replication in mammalian cells. Next, we tested whether these increases translated to higher virulence in mice. Animals were infected intranasally with different doses of wild-type or mutant TY93/H5N1 viruses and observed for survival to determine the amount of virus required to kill 50% of the infected animals (mouse lethal dose 50; MLD₅₀;

Fig. 5 and Supplementary Table 6). All mutant viruses (with the exception of TY93/H5N1 PB2-D740N) had a lower MLD₅₀ value than that of the wild-type virus, demonstrating increased virulence in mice compared with TY93/H5N1 virus. The greatest virulence in mice was detected for the PB1-N105S and PB2-E192K mutants, for which infection with one plaque-forming unit (p.f.u.) resulted in such severe clinical symptoms that all (for PB1-N105S) or two (PB2-E192K) of the infected mice had to be euthanized. The PB1-N105S mutant did not significantly increase the TY93/H5N1 polymerase activity in minireplicon assays (Fig. 4), again suggesting that the polymerase mutations influence steps beyond vRNA replication and transcription. We also found that the newly identified PB2-627V and -701V mutations had MLD₅₀ values identical to those of the known PB2-627K and -701N mutants, respectively (Supplementary Table 6). Mice infected with these viruses succumbed rapidly to infection (Fig. 5).

To assess virus titres in infected mice, we intranasally inoculated six mice each with 10⁴ p.f.u. of wild-type or mutant virus. Organs were collected from three mice each on days 3 and 5 post infection, except for mice infected with TY93/H5N1 PB2-E627K (all three mice were dead by day 5) or infected with PB2-D701V (2 of 3 infected mice were dead by day 5). Virus titres in the organs of infected mice were assessed by use of plaque assays in MDCK cells. On day 3 post infection, several mutant viruses replicated to higher titres than wild-type virus in the nasal turbinates and/or lungs of infected mice (Fig. 6 and Table 1), consistent with increased replicative ability in virus-infected cells (Fig. 3) and/or in minireplicons (Fig. 4). Some of these mutants were also detected in the brains of infected mice (Fig. 6). On day 5 post infection, the titres of wild-type and most mutant viruses were comparable in the lungs and nasal turbinates (Fig. 6 and Table 1). The PB2-740N mutant was not isolated from nasal turbinates, even though it replicated efficiently in the lungs of infected mice. Virus titres in the brains of infected mice differed greatly on day 5 post infection, ranging from undetectable to >10⁶ p.f.u. per g for mice infected with PB2-701N, and >10⁷ p.f.u. per g for a mouse infected with PB2-701V. Collectively, these data demonstrate that some of the mutations identified through our high-throughput screening approach increased the virulence of an avian H5N1 influenza virus in mice.

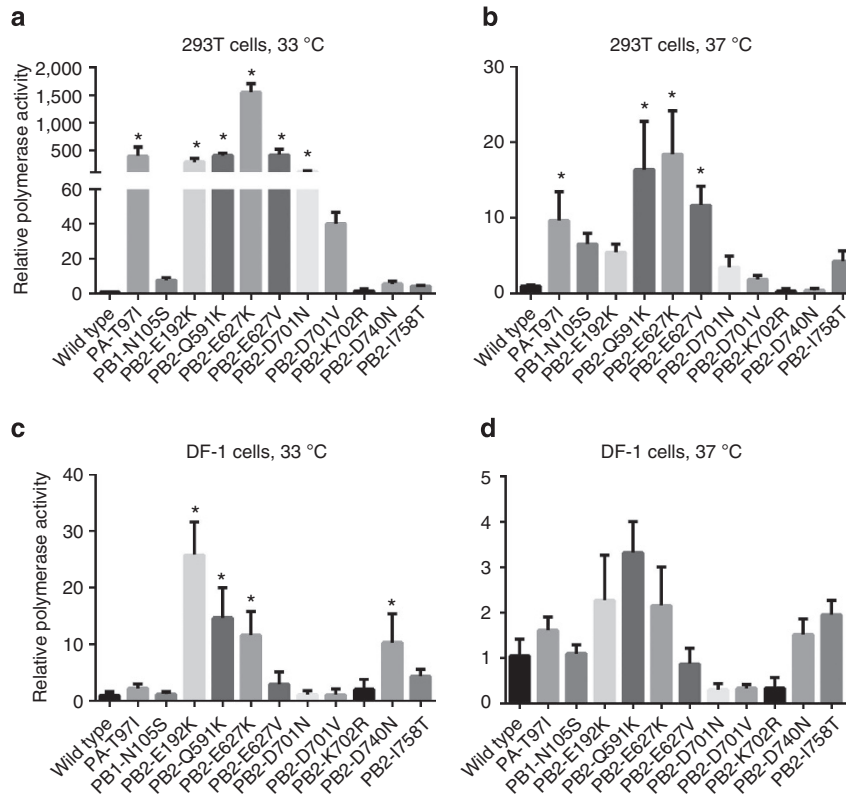


Figure 4 | Polymerase activity of wild type and polymerase mutants. Human 293 or avian DF-1 cells were transfected with plasmids expressing the wild type or mutant polymerase and NP proteins, and with a plasmid for the synthesis of a virus-like RNA encoding luciferase. In addition, cells were co-transfected with a plasmid encoding an internal control, Renilla luciferase/thymidine kinase (Promega). Cells were incubated for 48 h at 33 °C (a,c) or 37 °C (b,d), and then assessed in a dual-luciferase assay; the luciferase activity was normalized to the internal Renilla luciferase control. Data shown are the mean ± s.d. of three separate independent experiments and normalized to wild type ($n = 3 \pm \text{s.d.}$). The asterisk indicates a P value of < 0.05 compared with the polymerase activity of the wild-type polymerase complex (Kruskal-Wallis test).

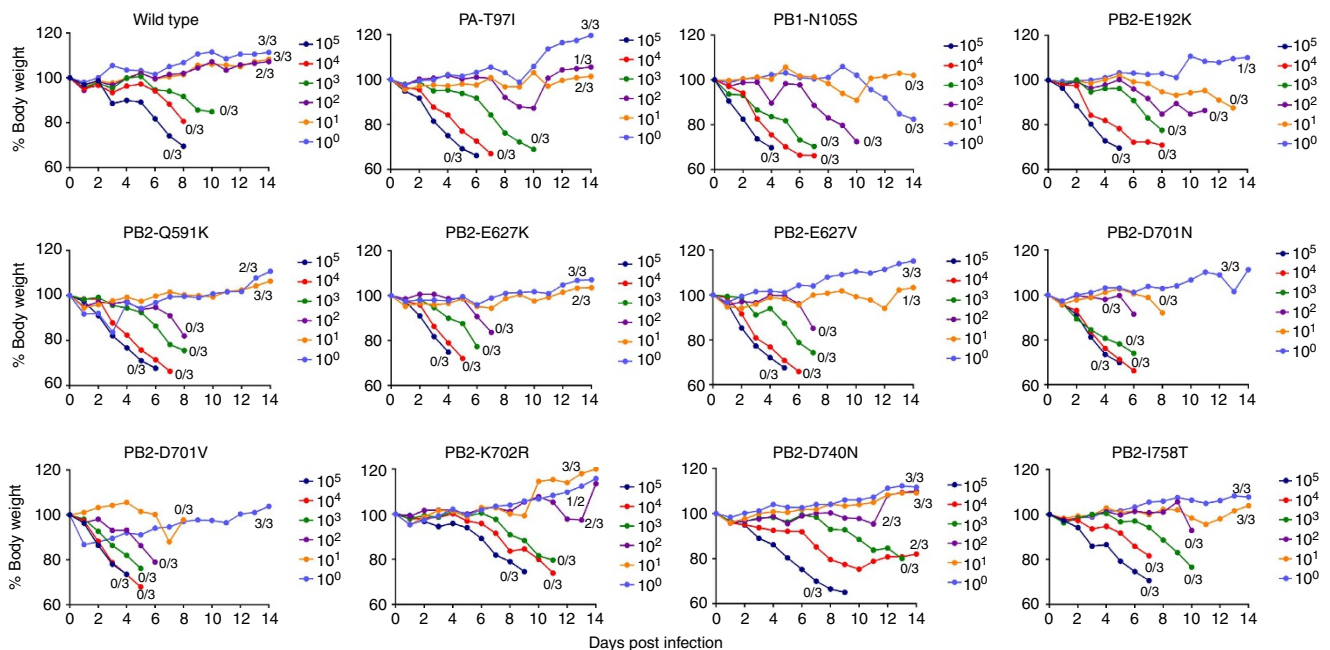


Figure 5 | Body weight changes and survival of mice infected with wild type or mutant TY93/H5N1 viruses. BALB/cByJ mice were infected with different doses ($1-10^5$ p.f.u.) of wild-type virus or the indicated polymerase mutant. Body weight was monitored daily; mice that lost $> 35\%$ of their starting weight or showed signs of severe disease were euthanized. Numbers shown indicate the number of surviving animals.

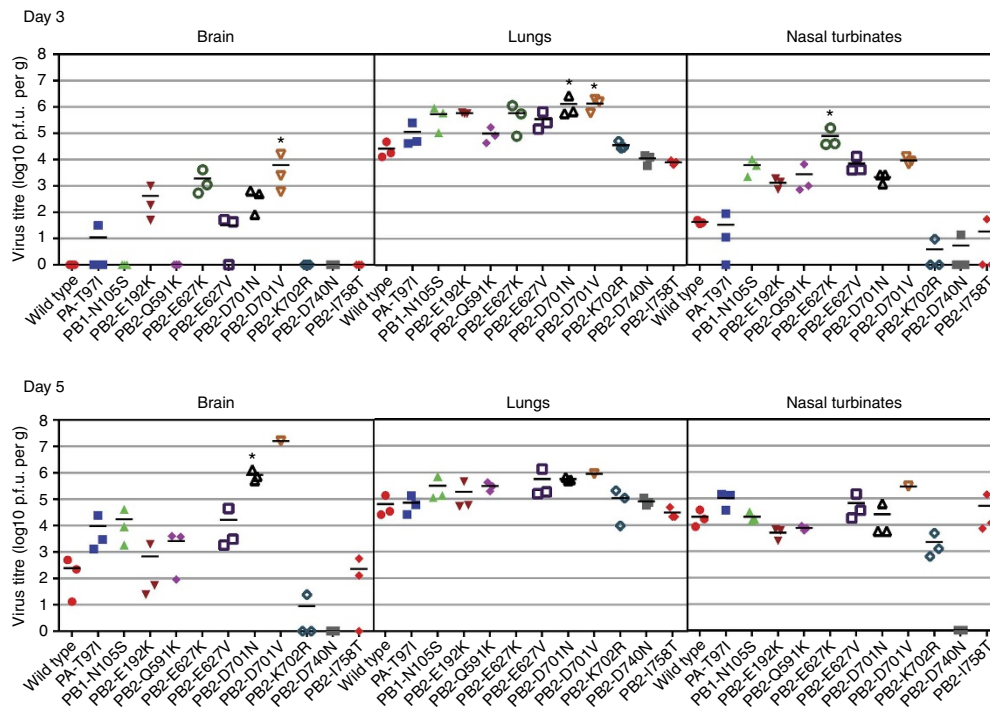


Figure 6 | Titres of wild-type and mutant TY93/H5N1 viruses in mouse organs. Mice were infected intranasally with 10^4 p.f.u. of the indicated virus. On days 3 and 5 post infection, organs were collected for virus titration in MDCK cells. Horizontal bars show the mean titres; asterisks indicate virus titres significantly different from that of wild-type TY93/H5N1 virus ($P < 0.05$; Dunnett's test). Note that 2 of 3 mice infected with TY93/H5N1 PB2-D701V died before day 5; hence, statistical analysis of data obtained for the remaining animal was not possible.

Discussion

By using a high-throughput screening approach, we identified several mutations in an avian H5N1 influenza virus polymerase complex that increase the replicative ability and virulence of the virus, and may facilitate avian influenza virus adaptation to mammals. Previous studies to identify mutations that promote mammalian adaptation and increase the virulence of avian influenza viruses have typically relied on the experimental and/or computational identification of signature amino-acid mutations found in natural isolates^{28–30}, or on sequential virus passages in mice^{24,31,32}, ferrets³³ or swine³⁴ (an approach known to increase virulence). However, sequential passages of avian H5N1 influenza viruses in mammals often lead to the selection of the dominant PB2-E627K (refs 35–37) mutation, which may mask the emergence of other adaptive mutations. The identification of host-adaptive mutations through sequence analysis is typically challenging because the observed mutations may be random and not associated with any adaptive advantage. To circumvent these potential drawbacks, we utilized a random mutagenesis approach to create a large pool of diverse mutants, followed by a screening process for variants that replicate efficiently in mammalian cells. This approach led us to identify specific mutations in the polymerase complex that, to the best of our knowledge, have never been reported as mammalian-adapting mutations, namely PB2-E192K, -E627V, -D701V, -K702R and PB1-N105S.

Glutamic acid is typically found at position 627 of PB2 in avian influenza viruses. Most human influenza viruses, with the exception of the 2009 pandemic H1N1 viruses, encode lysine at this position. Replication of avian influenza viruses in mice or humans frequently leads to the emergence of the PB2-E627K mutation^{26,38–41}. Lysine at position 627 is part of a basic patch at the surface of PB2, which is disrupted by a glutamic acid residue at this position^{42,43}. This patch may be essential for the interaction of PB2 with host factors. The valine residue at

position 627 of PB2 would result in a slightly basic surface patch. The PB2-627V mutation increased viral replication in mammalian cells and virulence in mice compared with PB2-627E; in fact, the replicative ability and virulence of TY93/H5N1 encoding PB2-627V were comparable to that of the virus encoding the mammalian-adapting PB2-627K residue. A recent study found that the PB2-627V mutation slightly reduced replicative ability *in vitro* compared with PB2-627K (ref. 44); however, this study was carried out in the background of the cell culture-adapted A/Puerto Rico/8/34 (H1N1) virus. Although we found that TY93/H5N1 viruses encoding PB2-627K or -627V have comparable replicative ability and mouse virulence, the PB2-627V variant is rarely found in natural influenza isolates. To date, it has only been detected in a single H7N9 human isolate (A/Hong Kong/5731/2014), five H1N1 swine isolates, a single H7N1 avian isolate (A/Pekin robin/California/30412/1994) and in several H9N2 isolates from poultry farms in Egypt and Israel between 2000 and 2013; ref. 45). We speculate that the type of nucleotide mutation may contribute to the selection of PB2-E627K over PB2-E627V: glutamic acid can mutate to lysine with a transition mutation in the first position (G to A); by contrast, the conversion of glutamic acid to valine requires a transversion mutation (GAG to GTG). Transition mutations are (de)aminations, whereas transversions require an exchange between a pyrimidine and a purine structure. Consequently, the transition:transversion ratio is reported to be $\sim 10:1$ (refs 46,47). This result highlights an advantage of our approach: by starting with error-prone PCR-generated virus libraries possessing random mutations, we were able to identify a novel putative mammalian-adapting mutation that had not been identified through surveillance or virus passages, likely due to its rare occurrence.

At position 701 of PB2, an aspartic acid (primarily found among avian influenza viruses) or asparagine residue (frequently detected

upon adaptation of avian influenza viruses to mammals^{17,18}) affects influenza virus host range and virulence, primarily through differences in their interaction with mammalian α -importins^{27,48}. Here we found that the replicative ability and virulence of TY93/H5N1 encoding PB2-701V is comparable to that of the PB2-701N variant. However, PB2-701V has been found in only one avian (A/duck/Zhejiang/0224-6/2011, H1N2; ref. 49) and one swine influenza virus (A/swine/Moegligen/IDT14859/2012, H1N2, KC631909). Similar to the PB2-E627K and -627V mutations discussed earlier, the commonly found PB2-D701N mutation can be achieved through a transition mutation, whereas the PB2-D701V mutation requires a transversion, possibly explaining why it is rarely found among natural isolates.

Based on computational analyses, position 702 in the PB2 protein has been suggested as a host marker^{50–52}. Most human influenza viruses encode arginine at this position, whereas avian influenza viruses typically possess lysine. The finding of seven human H5N1 viruses that encode PB2-702R (AF258843, AF258845, EU146728, CY014356, CY014359, CY014360 and AF258849) may further suggest a mammalian-adapting function for this residue. Our experimental data support this concept: TY93/H5N1 encoding PB2-702R replicated efficiently in mammalian cells and in the lungs of infected animals, and its MLD₅₀ value was lower than that of wild-type TY93/H5N1 virus (encoding PB2-702K). Given that this residue is located adjacent to the known mammalian-adapting marker at position 701 (see previous paragraph), the amino acid at position 702 may affect virulence and host range by modulating the interaction of PB2 with α -importins^{27,48}.

In our study, TY93/H5N1 virus possessing the PB2-E192K mutation was more virulent in mice than in wild-type TY93/H5N1 virus, and conferred efficient replication in mammalian cells. This mutation has only been found in eight natural isolates: three human seasonal influenza viruses (CY088395, CY007154 and CY098748), an H7N7 isolate from a sea mammal (GU050553), an avian H5N2 virus (CY094636), two H3N3 swine isolates (AY619970 and AY619962) and a human H5N1 isolate (CY098748). The function of this amino-acid position is currently unknown.

The PA-T97I mutation identified here was also detected upon adaptation of a low pathogenic avian H5N2 influenza virus (A/aquatic bird/Korea/WI81/2005) to mice⁵³. The highly virulent, mouse-adapted variant possessed numerous amino-acid changes across several viral proteins; further studies revealed that the PA-T97I mutation played a critical role in the mouse adaptation of this virus⁵⁴. These data are consistent with our findings and ascribe a role to PA-97I in the adaptation of avian influenza viruses to mammals.

The other mutations identified here (that is, PB1-N105S, PB2-D740N and PB2-I758T) have been detected in small numbers of human and/or avian influenza viruses; however, the biological function of these residues is not yet known.

Interestingly, some of the mutations identified here (such as PB2-E192K) increased the replicative ability of the polymerase complex in minireplicon assays not only in mammalian but also in avian cells (Table 1); yet, these increases did not translate into increased virus titres in avian cells. This finding suggests that these mutations support the replication and transcription of the viral genome, but negatively affect other steps in the viral life cycle (for example, the nuclear export of viral ribonucleoprotein complexes).

The identification of putative mammalian-adapting mutations in avian influenza viruses is invaluable for risk assessment of newly emerging strains, although not all of the mammalian-adapting mutations identified here may have the same effects in other influenza viruses. In the current study, we developed and implemented a high-throughput screening process to identify mutations in the polymerase complex of an avian influenza virus

that confer high replication in mammalian cells. All mutations identified in our study have been previously found in natural isolates, indicating that they may (re)emerge in novel strains and perhaps facilitate virus adaptation to mammals.

Methods

Cells. MDCK cells were maintained in Eagle's minimal essential media (MEM) containing 5% newborn calf serum. MDCK cells stably expressing the HA protein derived from A/WSN/33 (H1N1, WSN) were established by transduction with a retroviral vector and maintained in 5% newborn calf serum/MEM containing 800 $\mu\text{g ml}^{-1}$ Geneticin (Gibco). Chicken fibroblast (DF-1) and human embryonic fibroblast (293 and 293T) cells were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum (FCS). 293 Cells stably expressing the WSN HA protein were established by transduction with a retroviral vector and maintained in 10% FCS/MEM supplemented with 2 $\mu\text{g ml}^{-1}$ puromycin (Gibco). Adenocarcinomic human alveolar basal epithelial (A549) cells were maintained in a 1:1 mixture of DMEM and Ham's F12 nutrient medium (DF12; Invitrogen) supplemented with 10% FCS. Lung adenocarcinoma (Calu-3) cells were maintained in DMEM containing 10% FCS. All cells were maintained at 37 °C and 5% CO₂ with an antibiotic/antimycotic. 293, DF-1 and A549 cells were obtained from American Type Culture Collection (ATCC). Calu-3 cells from Dr Raymond Pickles (University of North Carolina-Chapel Hill), MDCK cells from Dr Robert Webster (St Jude Children's Research Hospital, Memphis, Tennessee) and 293T cells from Dr Tadashi Matsuda (Kansai Medical University, Japan).

Plasmid library construction. Random mutations were introduced into the targeted regions of the polymerase gene segments by PCR-based random mutagenesis using the Genomorph II kit (Stratagene). The targeted mutation rate (1–2 amino-acid changes per amplified region) was achieved by optimizing the quantity of the template, and was confirmed by sequence analysis of 24 randomly selected PCR products per reaction. The mutagenized regions were then cloned into the RNA polymerase I vectors containing the rest of the respective gene segment. To increase library titres, we constructed four libraries for each gene segment (Fig. 1c).

Reverse genetics and virus propagation. Replication-incompetent viruses were generated by using the reverse genetics system as described previously⁵⁵. Briefly, the eukaryotic protein expression plasmids encoding the WSN HA, NP, PA, PB1 and PB2 proteins under the control of the chicken β -actin promoter were transfected into 293T cells (Fig. 1b). The WSN virus was obtained from Dr Robert Webster. Cells were co-transfected with RNA polymerase I plasmids transcribing the viral M, NA and NS segments of WSN virus and the viral PB2, PB1, PA and NP segments of A/Muscovy duck/VN/TY93/2007 (H5N1) virus obtained from Dr Mai Le (National Institute of Hygiene and Epidemiology, Hanoi, Vietnam; Fig. 1b). Moreover, cells were transfected with the pPolIIHA(48)GFP(291) plasmid, which encodes a virus-like RNA possessing the 3' influenza viral promoter and noncoding regions, 48 nucleotides that correspond to the HA coding sequence at the 3' end of the vRNA (maintained because of their role in vRNA packaging⁵⁶), the GFP coding sequence, 291 nucleotides that correspond to the HA coding sequence at the 5' end of the vRNA (maintained because of their role in vRNA packaging⁵⁶), and finally the 5' HA noncoding and viral promoter region (Fig. 1a). The mutant virus libraries were generated similarly but we replaced the wild-type pPolII plasmid for PB2, PB1 or PA with the corresponding mutant plasmid library. Mutant virus library stocks were generated in 293-HA cells, and titres were determined by plaque assay in MDCK-HA cells.

Selected mutations were also tested in the background of a reassortant virus possessing the PB2, PB1, PA and NP vRNA segments of TY93/H5N1 virus and the HA, NA, M and NS vRNA segments of H5N1 A/Vietnam/1203/2004 virus (obtained from Dr Mai Le) or in the background of authentic TY93/H5N1 virus. Virus stocks were generated in MDCK cells or specific pathogen-free embryonated chicken eggs and stored at –80 °C until use. All cell culture experiments with wild-type and mutant TY93/H5N1 viruses were carried out in BSL3 + containment; animal experiments were performed in BSL3-Ag containment.

Fluorescence-activated cell sorting. 293-HA cells were infected with the mutant virus libraries and wild-type controls at an MOI of 0.1. At 1 h post infection, virus supernatant was removed and replaced with 0.3% bovine serum albumin (BSA)/MEM infection media. At 4 h post infection, cells were treated with 0.25% trypsin-EDTA and resuspended in 0.3% BSA/MEM. FACS was carried out using a FACSAria Cell Sorter (BD Biosciences) equipped with the FACSDiva software suite (BD Biosciences). A baseline fluorescence level was determined by analysing 10,000 cells infected with parental TY93/H5N1 GFP-627E virus. The cell sorting gate was then set to capture cells infected with a mutant virus library that expressed GFP at higher intensity than that observed for TY93/H5N1 GFP-627E. The cell sorting gate was adjusted every hour by analysing 10,000 cells infected with TY93/H5N1 GFP-627E virus. Single cells with increased GFP expression levels were sorted into 96-well plates seeded with 1.5×10^5 MDCK-HA cells per well and incubated at 37 °C. At 48 or 72 h post infection, GFP levels were measured on an Infini MH1000 plate reader (Tecan). Supernatants from GFP-positive wells were

transferred to another 96-well plate by using a MICROLAB STAR liquid handling workstation (Hamilton), aseptically sealed with a PlateLoc Thermal Microplate Sealer (Agilent) and stored at -80°C .

Quantitative RT-PCR. 293 Cells were seeded at a density of 4×10^5 cells per well. Twenty-four hours later, 50 μl of virus was added to each well. Six hours after infection, total RNA was isolated on a Biorobot Universal System (Qiagen) by using the RNeasy 96 kit (Qiagen). In parallel, 293 cells were infected with the same amount of virus and total RNA was extracted immediately ($t = 0$). qRT-PCR was performed on a 7900HT Fast Real-time PCR system (Applied Biosystems) by using the Superscript III One-Step Platinum qRT-PCR kit (Life Technologies) and the following cycling conditions: 37°C for 15 min, 50°C for 30 min and 50 cycles of 95°C for 2 min and 62°C for 1 min. Oligonucleotide primer and probe sequences were adapted from van Elden *et al.*⁵⁷ and are as follows: forward primer—5'-GGACTGCAGCGTAGACGCTTT-3', reverse primer—5'-CATCTGTTGTATATGAGGCCAT-3', probe—5'-6FAM-CTCAGTTATTCTGCTGGTG CACTTGCC-BHQ-3'. Relative quantification of the change in M levels between 0 and 6 h post infection was performed by using the $\Delta\Delta\text{Ct}$ method⁵⁸.

Deep sequencing. Total RNA was isolated as described in the previous section. The NP and polymerase genes were amplified using a nested PCR approach. First, a multisegmented RT-PCR comprising 25 cycles was performed for each sample. Next, nested PCR reactions comprised of 28 cycles were performed to generate two overlapping amplicons for each polymerase segment and a single amplicon for the NP segment. All primer sequences are provided in Supplementary Table 7. Nested amplicons were then combined into an amplicon pool, which was purified by using AMPure Beads (Beckman Coulter). The purified amplicon pool was processed with the Ion Xpress Plus Fragment Library Kit (Life Technologies) to generate sheared, barcoded and size-selected Ion Torrent PGM 200-base libraries, and then amplified in eight additional PCR cycles to increase the number of barcoded library fragments. Final pools for sequence analysis were constructed by combining equimolar amounts of the individually barcoded libraries (which correspond to a single isolate). The pools were then sequenced on the Ion Torrent PGM with 314 chips. Basecalling was performed by using TraceTuner⁵⁹, and the resulting sequences were quality-trimmed and then assembled using the merger utility from the EMBOSS software suite⁶⁰. All processing steps were integrated into a reproducible Galaxy⁶¹ workflow and are available upon request. The resulting assembled sequences were analysed by using the Galaxy bioinformatics suite⁶¹ and the Integrative Genomics Viewer⁶².

Viral growth kinetics. Cell infections were performed in triplicate in six-well Biocoat poly-D-lysine-coated plates (Becton Dickinson). Cells, seeded at a density of 3×10^5 cells per well 24 h before infection, were infected with individual viruses at an MOI of 0.001. Viruses were adsorbed to cells for 1 h at 37°C in MEM/0.3% BSA medium. Then, cells were washed with PBS and overlaid with MEM/0.3% BSA containing $1 \mu\text{g ml}^{-1}$ of tosyl phenylalanyl chloromethyl ketone (TPCK) treated trypsin for human 293, A549 and Calu3 cells; avian DF-1 cells were incubated without trypsin. Supernatants were collected at 24, 48 and 72 h post infection, and plaque assays were carried out in MDCK cells to determine virus titres.

Minireplicon assays. The polymerase activity of viral ribonucleoprotein (RNP) complexes was measured by using a dual luciferase reporter system (Promega). A plasmid transcribing a virus-like RNA encoding the luciferase reporter gene (pPOLWSNNA F-Luc) was transfected into 6×10^4 293T cells together with protein expression plasmids for wild-type or mutant TY93/H5N1 PB2, PB1, PA and NP proteins. For minireplicon assays in chicken DF-1 cells, pPOLWSNNA F-Luc was replaced with pCk-F-Luc, which possesses an avian RNA polymerase I promoter. Cells were co-transfected with a protein expression plasmid for Renilla luciferase/thymidine-kinase (Promega), which served as an internal control to measure transfection efficiency. Cells were incubated at 33 or 37°C for 48 h, and luciferase activity was measured on an Infinite M1000 plate reader (Tecan). Luciferase activity was normalized to the internal Renilla luciferase control.

Mouse studies. MLD₅₀ values and tissue titres of wild-type and mutant viruses were determined in 6-week-old female BALB/cByJ mice (The Jackson Laboratory). For MLD₅₀ determination, three mice per group were anaesthetized by intraperitoneal injection with ketamine and dexmedetomidine, and intranasally inoculated with 100 μl of different virus doses ($1-10^5$ p.f.u.) diluted in MEM/0.3% BSA. After inoculation, anaesthesia was reversed with an intraperitoneal injection of atipamazole. Body weight and survival status were monitored daily. Mice exhibiting the following clinical signs (rapid or slow breathing, ruffled fur, rapid weight loss, hunched posture, shivering, inappetence, diarrhoea or constipation, lethargy or any obvious illnesses such as superficial skin injury) singly or in combination were euthanized if the clinical signs did not improve within 48 h, and resolution was determined to be improbable. In addition, mice that lost $>35\%$ of their starting body weight or showed signs of the inability to remain upright singly or in

combination with any of the other clinical signs listed above were euthanized. MLD₅₀ values were calculated by using the method of Reed and Muench⁶³.

Virus titres in organs were determined by intranasally infecting six mice per group with 10^4 p.f.u. of virus. At 3 or 5 days post infection, mice were euthanized, and brain, lung and nasal turbinates were isolated and stored at -80°C . All mice infected with TY93/H5N1 PB2-E627K, and 2 of 3 mice infected with TY93/H5N1 PB2-D701V died before day 5 and were excluded from further analyses. Organs were homogenized in MEM/0.3% BSA by using a TissueLyser (Qiagen) and serially diluted in MEM/0.3% BSA. Virus titres were measured by performing plaque assays in MDCK cells. Animal studies were performed in accordance with the University of Wisconsin-Madison Animal Care and Use committee protocols.

Statistical analysis. The data were analysed by using the R software (www.r-project.org), version 3.1, and Prism (GraphPad Software Inc.). For comparisons of measurements from multiple groups collected at a single time point (that is, minireplicon data), we used the Kruskal-Wallis test followed by pairwise comparisons of the groups and appropriate adjustments of significance. For comparisons of multiple groups with dependent measurements (that is, viral growth curves in cell culture for which aliquots were collected from the same culture at different time points), we log-transformed the data and fitted a linear mixed-effects model to it. Next, we built a contrast matrix to compare the strains in a pairwise fashion at the same time points. The P values were adjusted by using Holm's method to account for multiple comparisons. Mouse survival data were analysed by using the log-rank (Mantel-Cox) test. Results were considered statistically significant for P values (or adjusted P values) of <0.05 .

Biosafety and biosecurity. These studies were conducted after the University of Wisconsin-Madison (UW) Select Agent Program completed risk assessments and the IBC approved the experimental protocols. Research updates were submitted to the Alternate Responsible Official of the University of Wisconsin-Madison Select Agent Program. In addition, risk mitigation plans approved by National Institutes of Health (NIH) for NIH-funded research grants that contain Dual Use Research of Concern DURC were followed.

The manuscript was submitted to National Institute of Allergy and Infectious Diseases (NIAID), who concluded that the research constitutes DURC, and that the manuscript should be reviewed by the local IBC. In parallel, the manuscript was submitted to the University of Wisconsin-Madison DURC Subcommittee and their findings were presented to the University of Wisconsin-Madison IBC and the University of Wisconsin-Madison Biosecurity Task Force (BTF). The University of Wisconsin-Madison BTF regularly reviews the research programme and ongoing laboratory activities. The task force has a diverse skill set and provides support in the areas of biosafety, facilities, research compliance, security, law, healthcare and public health. Members of the BTF are in frequent contact with the Select Agent Program, the principal investigators and laboratory personnel to provide oversight and assure biosecurity. All three panels (that is, the UW-IBC, UW-DURC Subcommittee and UW-BTF) concluded that this study constitutes DURC, but that proper risk mitigation strategies were followed and that the data should be published in full.

All mouse studies with H5N1 viruses were performed in BSL3Ag containment laboratories by two experienced PhD-level scientists. Staff working in BSL3Ag wear disposable overalls and powered air-purifying respirators that filter the air, and shower out on exit from the facility. The BSL3Ag facility at the University of Wisconsin-Madison was designed to exceed the standards outlined in Biosafety in Microbiological and Biomedical Laboratories (5th edition; <http://www.cdc.gov/biosafety/publications/bml5/BMBL.pdf>). Features include controlled access, entry/exit through a shower change room, effluent decontamination, negative air-pressure, double-door autoclaves, gas decontamination ports, plus HEPA-filtered supply and double-HEPA-filtered exhaust air, double-gasketed watertight and airtight seals, airtight dampers on all ductwork, and the structure is pressure-decay tested regularly. The University of Wisconsin-Madison facility has a dedicated alarm system that monitors all building controls and sends alarms (~ 500 possible alerts). Redundancies and emergency resources are built-in to the facility including two air handlers, two compressors, two filters in each location where filters are needed, two effluent sterilization tanks, two power feeds to the building, an emergency generator in case of a power failure and other physical containment measures in the facility that operate without power. Biosecurity monitoring of the facility is ongoing. All personnel undergo Select Agent security risk assessment by the United States Criminal Justice Information Services Division and complete rigorous biosafety, BSL3 and Select Agent training before participating in BSL3-level experiments. Refresher training, including drills and review of emergency plans, is scheduled on a regular basis. The principal investigator participates in training sessions and emphasizes compliance to maintain safe operations and a responsible research environment. The laboratory occupational health plan is in compliance with the University of Wisconsin-Madison Occupational Medicine Program. Select Agent virus inventory, secured behind two physical barriers, is checked each month and documentation is submitted to the University of Wisconsin-Madison Select Agent Program Manager. Updates on virus inventory (which is carried out every month) are submitted 1–2 times per year to the file holder in the Select Agent branch of the Centers for Disease Control and Prevention (CDC). The research programme, procedures, occupational health plan,

documentation, security and facilities are reviewed annually by the University of Wisconsin-Madison Responsible Official and at regular intervals by the CDC, and the Animal and Plant Health Inspection Service as part of the University of Wisconsin-Madison Select Agent Program.

All experiments were completed before 17 October 2014, when the United States Government announced a voluntary pause on 'certain gain-of-function experiments involving influenza, SARS and MERS viruses'.

Ethics. Our experiments in mice followed the University of Wisconsin-Madison's Animal Care and Use Protocol. All experiments were approved by the Animal Care and Use Committee of the University of Wisconsin-Madison (protocol number V00806), which acknowledged and accepted both the legal and ethical responsibility for the animals, as specified in the Fundamental Guidelines for Proper Conduct of Animal Experiment and Related Activities in the Animal Welfare Act and associated Animal Welfare Regulations and Public Health Service Policy (USA).

References

- Butt, K. M. *et al.* Human infection with an avian H9N2 influenza A virus in Hong Kong in 2003. *J. Clin. Microbiol.* **43**, 5760–5767 (2005).
- Shi, W., Shi, Y., Wu, Y., Liu, D. & Gao, G. F. Origin and molecular characterization of the human-infecting H6N1 influenza virus in Taiwan. *Protein Cell* **4**, 846–853 (2013).
- Fouchier, R. A. *et al.* Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. *Proc. Natl Acad. Sci. USA* **101**, 1356–1361 (2004).
- Qi, W. *et al.* Genesis of the novel human-infecting influenza A(H10N8) virus and potential genetic diversity of the virus in poultry, China. *Euro Surveill.* **19** (2014).
- Ostrowsky, B. *et al.* Low pathogenic avian influenza A (H7N2) virus infection in immunocompromised adult, New York, USA, 2003. *Emerg. Infect. Dis.* **18**, 1128–1131 (2012).
- Tweed, S. A. *et al.* Human illness from avian influenza H7N3, British Columbia. *Emerg. Infect. Dis.* **10**, 2196–2199 (2004).
- Imai, M. *et al.* Transmission of influenza A/H5N1 viruses in mammals. *Virus Res.* **178**, 15–20 (2013).
- Knipe, D. M. & Howley, P. M. *Fields Virology* 6th edn. (Wolters Kluwer/Lippincott Williams & Wilkins Health, 2013).
- Cauldwell, A. V., Long, J. S., Moncorge, O. & Barclay, W. S. Viral determinants of influenza A virus host range. *J. Gen. Virol.* **95**, 1193–1210 (2014).
- Matrosovich, M. *et al.* 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 (2000).
- Rogers, G. N. & Paulson, J. C. Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. *Virology* **127**, 361–373 (1983).
- Rogers, G. N. *et al.* Single amino acid substitutions in influenza haemagglutinin change receptor binding specificity. *Nature* **304**, 76–78 (1983).
- Hatta, M., Gao, P., Halfmann, P. & Kawaoka, Y. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. *Science* **293**, 1840–1842 (2001).
- Subbarao, E. K., London, W. & Murphy, B. R. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *J. Virol.* **67**, 1761–1764 (1993).
- Yamada, S. *et al.* Biological and structural characterization of a host-adapting amino acid in influenza virus. *PLoS Pathog.* **6**, e1001034 (2010).
- Mehle, A. & Doudna, J. A. Adaptive strategies of the influenza virus polymerase for replication in humans. *Proc. Natl Acad. Sci. USA* **106**, 21312–21316 (2009).
- Li, Z. *et al.* Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. *J. Virol.* **79**, 12058–12064 (2005).
- Gabriel, G. *et al.* The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proc. Natl Acad. Sci. USA* **102**, 18590–18595 (2005).
- Fan, S. *et al.* Novel residues in avian influenza virus PB2 protein affect virulence in mammalian hosts. *Nat. Commun.* **5**, 5021 (2014).
- Bussey, K. A., Bousse, T. L., Desmet, E. A., Kim, B. & Takimoto, T. PB2 residue 271 plays a key role in enhanced polymerase activity of influenza A viruses in mammalian host cells. *J. Virol.* **84**, 4395–4406 (2010).
- Zhou, B. *et al.* PB2 residue 158 is a pathogenic determinant of pandemic H1N1 and H5 influenza A viruses in mice. *J. Virol.* **85**, 357–365 (2011).
- Cheng, K. *et al.* PB2-E627K and PA-T97I substitutions enhance polymerase activity and confer a virulent phenotype to an H6N1 avian influenza virus in mice. *Virology* **468–470**, 207–213 (2014).
- Martinez-Sobrido, L. *et al.* Hemagglutinin-pseudotyped green fluorescent protein-expressing influenza viruses for the detection of influenza virus neutralizing antibodies. *J. Virol.* **84**, 2157–2163 (2010).
- Li, Z. *et al.* Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. *J. Virol.* **79**, 12058–12064 (2005).
- Li, Z. *et al.* Mutational analysis of conserved amino acids in the influenza A virus nucleoprotein. *J. Virol.* **83**, 4153–4162 (2009).
- Hatta, M. *et al.* Growth of H5N1 influenza A viruses in the upper respiratory tracts of mice. *PLoS Pathog.* **3**, 1374–1379 (2007).
- Gabriel, G., Herwig, A. & Klenk, H. D. Interaction of polymerase subunit PB2 and NP with importin alpha1 is a determinant of host range of influenza A virus. *PLoS Pathog.* **4**, e11 (2008).
- Smith, G. J. *et al.* Evolution and adaptation of H5N1 influenza virus in avian and human hosts in Indonesia and Vietnam. *Virology* **350**, 258–268 (2006).
- Subbarao, K. *et al.* Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. *Science* **279**, 393–396 (1998).
- Li, Y. G. *et al.* Characterization of H5N1 influenza viruses isolated from humans in vitro. *Virol. J.* **7**, 112 (2010).
- Li, Q. *et al.* Adaptation of a natural reassortant H5N2 avian influenza virus in mice. *Vet. Microbiol.* **172**, 568–574 (2014).
- Ping, J. *et al.* Genomic and protein structural maps of adaptive evolution of human influenza A virus to increased virulence in the mouse. *PLoS ONE* **6**, e21740 (2011).
- Herfst, S. *et al.* Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* **336**, 1534–1541 (2012).
- Wei, K. *et al.* Influenza A virus acquires enhanced pathogenicity and transmissibility after serial passages in swine. *J. Virol.* **88**, 11981–11994 (2014).
- Min, J. Y. *et al.* Mammalian adaptation in the PB2 gene of avian H5N1 influenza virus. *J. Virol.* **87**, 10884–10888 (2013).
- Hatta, M. *et al.* Growth of H5N1 influenza A viruses in the upper respiratory tracts of mice. *PLoS Pathog.* **3**, 1374–1379 (2007).
- Li, J. *et al.* Single mutation at the amino acid position 627 of PB2 that leads to increased virulence of an H5N1 avian influenza virus during adaptation in mice can be compensated by multiple mutations at other sites of PB2. *Virus Res.* **144**, 123–129 (2009).
- Min, J. Y. *et al.* Mammalian adaptation in the PB2 gene of avian H5N1 influenza virus. *J. Virol.* **87**, 10884–10888 (2013).
- Munster, V. J. *et al.* The molecular basis of the pathogenicity of the Dutch highly pathogenic human influenza A H7N7 viruses. *J. Infect. Dis.* **196**, 258–265 (2007).
- de Jong, R. M. *et al.* Rapid emergence of a virulent PB2 E627K variant during adaptation of highly pathogenic avian influenza H7N7 virus to mice. *Virol. J.* **10**, 276 (2013).
- Mase, M. *et al.* Recent H5N1 avian influenza A virus increases rapidly in virulence to mice after a single passage in mice. *J. Gen. Virol.* **87**, 3655–3659 (2006).
- Kuzuhara, T. *et al.* Structural basis of the influenza A virus RNA polymerase PB2 RNA-binding domain containing the pathogenicity-determinant lysine 627 residue. *J. Biol. Chem.* **284**, 6855–6860 (2009).
- Tarendeau, F. *et al.* Host determinant residue lysine 627 lies on the surface of a discrete, folded domain of influenza virus polymerase PB2 subunit. *PLoS Pathog.* **4**, e1000136 (2008).
- Chin, A. W. *et al.* Influenza A viruses with different amino acid residues at PB2-627 display distinct replication properties in vitro and in vivo: revealing the sequence plasticity of PB2-627 position. *Virology* **468–470C**, 545–555 (2014).
- Golender, N. *et al.* Genetic characterization of avian influenza viruses isolated in Israel during 2000–2006. *Virus genes* **37**, 289–297 (2008).
- Li, X. *et al.* Global and Local Persistence of Influenza A(H5N1) Virus. *Emerg. Infect. Dis.* **20**, 1287–1295 (2014).
- Weltman, J. K., Skowron, G. & Loriot, G. B. Influenza A H5N1 hemagglutinin cleavable signal sequence substitutions. *Biochem. Biophys. Res. Commun.* **352**, 177–180 (2007).
- Gabriel, G. *et al.* Differential use of importin-alpha isoforms governs cell tropism and host adaptation of influenza virus. *Nat. Commun.* **2**, 156 (2011).
- Wu, H. B. *et al.* Genetic characterization of subtype H1 avian influenza viruses isolated from live poultry markets in Zhejiang Province, China, in 2011. *Virus Genes* **44**, 441–449 (2012).
- Finkelstein, D. B. *et al.* Persistent host markers in pandemic and H5N1 influenza viruses. *J. Virol.* **81**, 10292–10299 (2007).
- Shaw, M. *et al.* Molecular changes associated with the transmission of avian influenza A H5N1 and H9N2 viruses to humans. *J. Med. Virol.* **66**, 107–114 (2002).
- Miotto, O. *et al.* Complete-proteome mapping of human influenza A adaptive mutations: implications for human transmissibility of zoonotic strains. *PLoS ONE* **5**, e9025 (2010).
- Song, M. S. *et al.* The polymerase acidic protein gene of influenza A virus contributes to pathogenicity in a mouse model. *J. Virol.* **83**, 12325–12335 (2009).
- Fan, S. *et al.* Two amino acid residues in the matrix protein M1 contribute to the virulence difference of H5N1 avian influenza viruses in mice. *Virology* **384**, 28–32 (2009).

55. Neumann, G. *et al.* Generation of influenza A viruses entirely from cloned cDNAs. *Proc. Natl Acad. Sci. USA* **96**, 9345–9350 (1999).
56. Fujii, Y., Goto, H., Watanabe, T., Yoshida, T. & Kawaoka, Y. Selective incorporation of influenza virus RNA segments into virions. *Proc. Natl Acad. Sci. USA* **100**, 2002–2007 (2003).
57. van Elden, L. J., Nijhuis, M., Schipper, P., Schuurman, R. & van Loon, A. M. Simultaneous detection of influenza viruses A and B using real-time quantitative PCR. *J. Clin. Microbiol.* **39**, 196–200 (2001).
58. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–408 (2001).
59. Denisov, G. A., Arehart, A. B. & Curtin, M. D. Data processing, computers, Patent US Patent 6,681,186 (2004).
60. Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* **16**, 276–277 (2000).
61. Blankenberg, D. *et al.* Galaxy: a web-based genome analysis tool for experimentalists. *Curr. Protoc. Mol. Biol.* **Chapter 19**, Unit 19.10. 1–21 (2010).
62. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
63. Reed, L. & Muench, H. A simple method of estimating fifty per cent endpoints. *Am. J. Hyg.* **27**, 493–497 (1939).

Acknowledgements

We thank Susan Watson for scientific editing and Julie A. Eggenberger, Peter J. Jester and Zachary Najacht for their technical assistance. This work was supported by a grant from the Bill & Melinda Gates Foundation (OPPGH5383), by the Japan Initiative for Global Research Network on Infectious Diseases from the Ministry of Education, Culture, Sports, Science, and Technology, Japan, by ERATO, Japan, by National Institutes of Health Office of Research and Infrastructure Programs grant OD 011106 to T.C.F. and

by a National Institute of Allergy and Infectious Diseases Public Health Service research grant under award number R01AI069274.

Author contributions

A.S.T., E.A.M., G.N. and Y.K. designed the experiments. A.S.T., P.J.H., L.H.-B., T.C.F. and M.H. performed the experiments. M.O. established HA-deficient influenza virus. A.F., J.V.D. and E.G. performed deep sequencing of virus samples. A.S.T. and C.A.M. analysed deep-sequencing data. T.J.S.L. performed statistical analyses. A.S.T., E.A.M., G.N. and Y.K. analysed all of the experimental data. A.S.T. and G.N. wrote the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Taft, A.S. *et al.* Identification of mammalian-adapting mutations in the polymerase complex of an avian H5N1 influenza virus. *Nat. Commun.* **6**:7491 doi: 10.1038/ncomms8491 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>