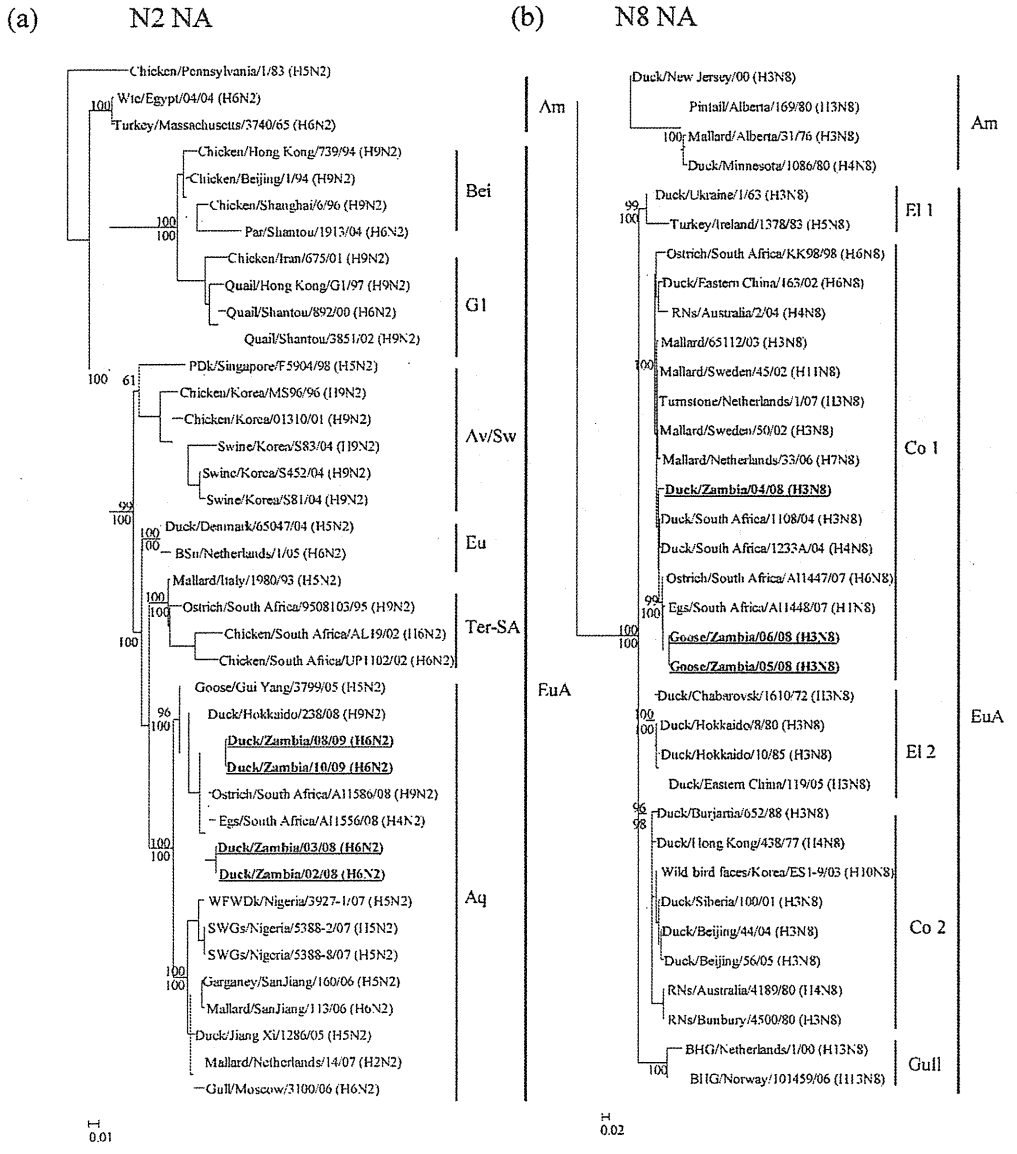

Fig. 2. Phylogenetic relationships of the N2 NA (a) and N8 NA (b) genes of AIVs isolated from wild birds in Zambia. Analysis was based on nt 78–1396 (1319 bp) of N2 NA and 54–1343 (1290 bp) of N8 NA. Numbers above and below branch nodes indicate neighbour-joining bootstrap values of ⩾50 % and Bayesian posterior probabilities of >95 %, respectively. Owing to space limitations, not all supports are indicated. Virus strains characterized in the present study are in bold and underlined. Bars, number of substitutions per site. Lineages: Aq, aquatic; Av/Sw, avian/swine; Co, contemporary; Ter-SA, terrestrial, South Africa. Strain names: BHG, black-headed gull; Egs, Egyptian goose; PDk, pekin duck; RNs, red-necked stint; SWGs, spur-winged goose; WFWDk, white-faced whistling duck. Other abbreviations are listed in the legend of Fig. 1.

Europe. The NP gene of Zb01 (H3N6) was closely related to that of ostrich/South Africa/AI1447/07 (H6N8) and both these strains belonged to a group of viruses comprising early and contemporary strains.

The matrix (M) gene tree showed that all the viruses isolated from wild birds in Zambia were in group 3, but they did not all cluster together (Fig. 4b). The majority of the viruses reported in this study grouped with isolates obtained from wild and domestic birds in South Africa and appear to have been derived from A/mallard/Netherlands/1/06 (H8N4)-like viruses. The M genes of Zb01 (H3N6) and Zb04 (H3N8) were closely related to that of turkey/Italy/3620/99 (H7N1), whilst those of Zb02 (H6N2), Zb03 (H6N2), Zb05 (H3N8) and Zb06 (H3N8) grouped with that of an H9N2 virus isolated from an ostrich in South Africa.

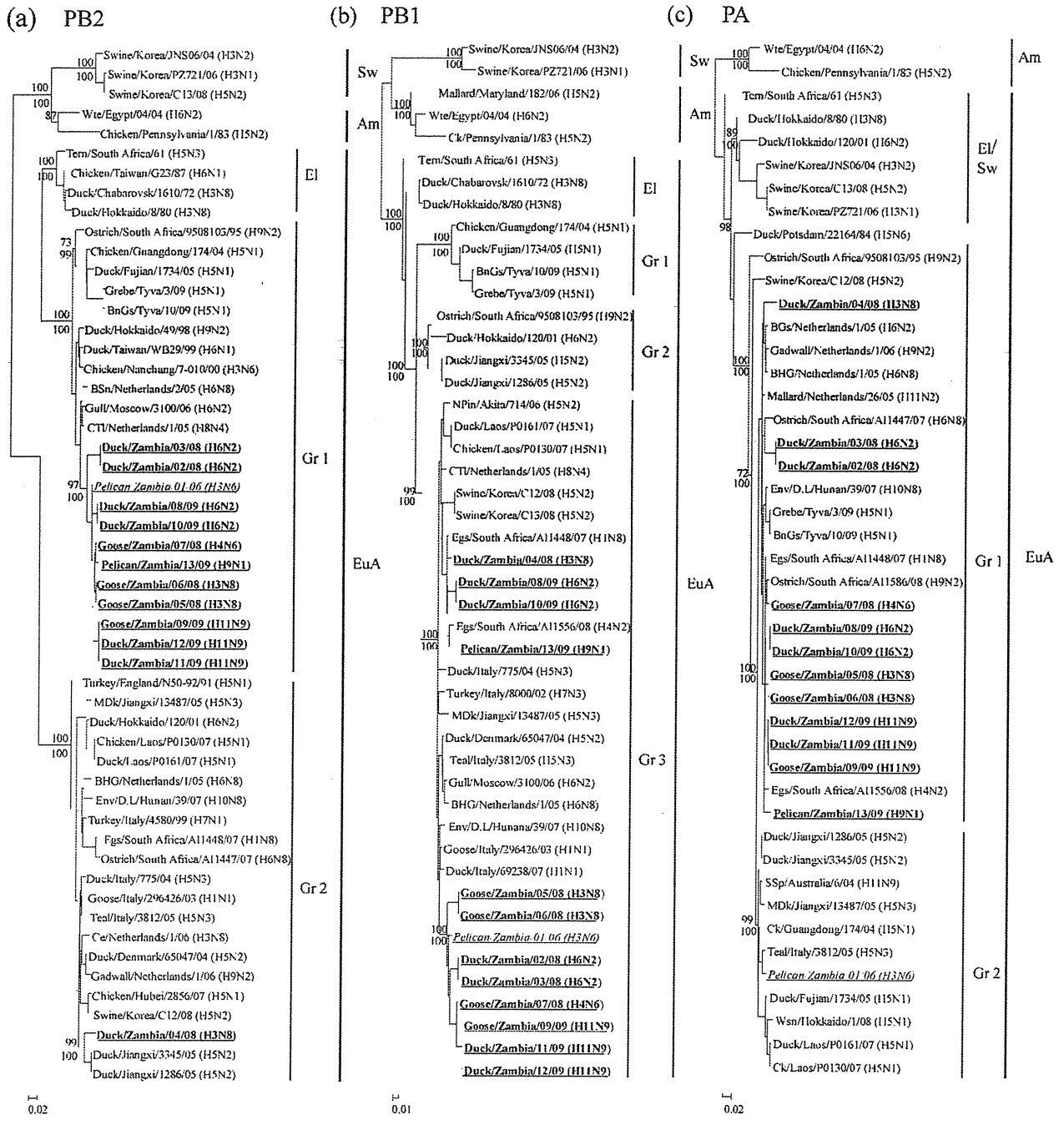

Fig. 3. Phylogenetic relationships of the PB2 (a), PB1 (b) and PA (c) genes of AIVs isolated from wild birds in Zambia. Analysis was based on nt 56–2285 (2230 bp) of PB2, 64–2281 (2218 bp) of PB1 and 30–2098 (2069 bp) of PA. Numbers above and below branch nodes indicate neighbour-joining bootstrap values of ⩾50 % and Bayesian posterior probabilities of >95 %, respectively. Due to space limitations, not all supports are indicated. Virus strains characterized in the present study are in bold and underlined, whilst the first AIV isolate in Zambia is italicized and underlined. Bars, number of substitutions per site. Lineages: El/Sw, early/swine; Sw, swine. Strain names: BGs, barnacle goose; BnGs, bean goose; Ce, common eider; Ck, chicken; CTl, common teal; D.L, Dongting Lake; Env, environment; MDk, migratory duck; SSp, sharp-tailed sandpiper; Wsn, whooper swan. Other abbreviations are listed in the legends of Figs 1 and 2.

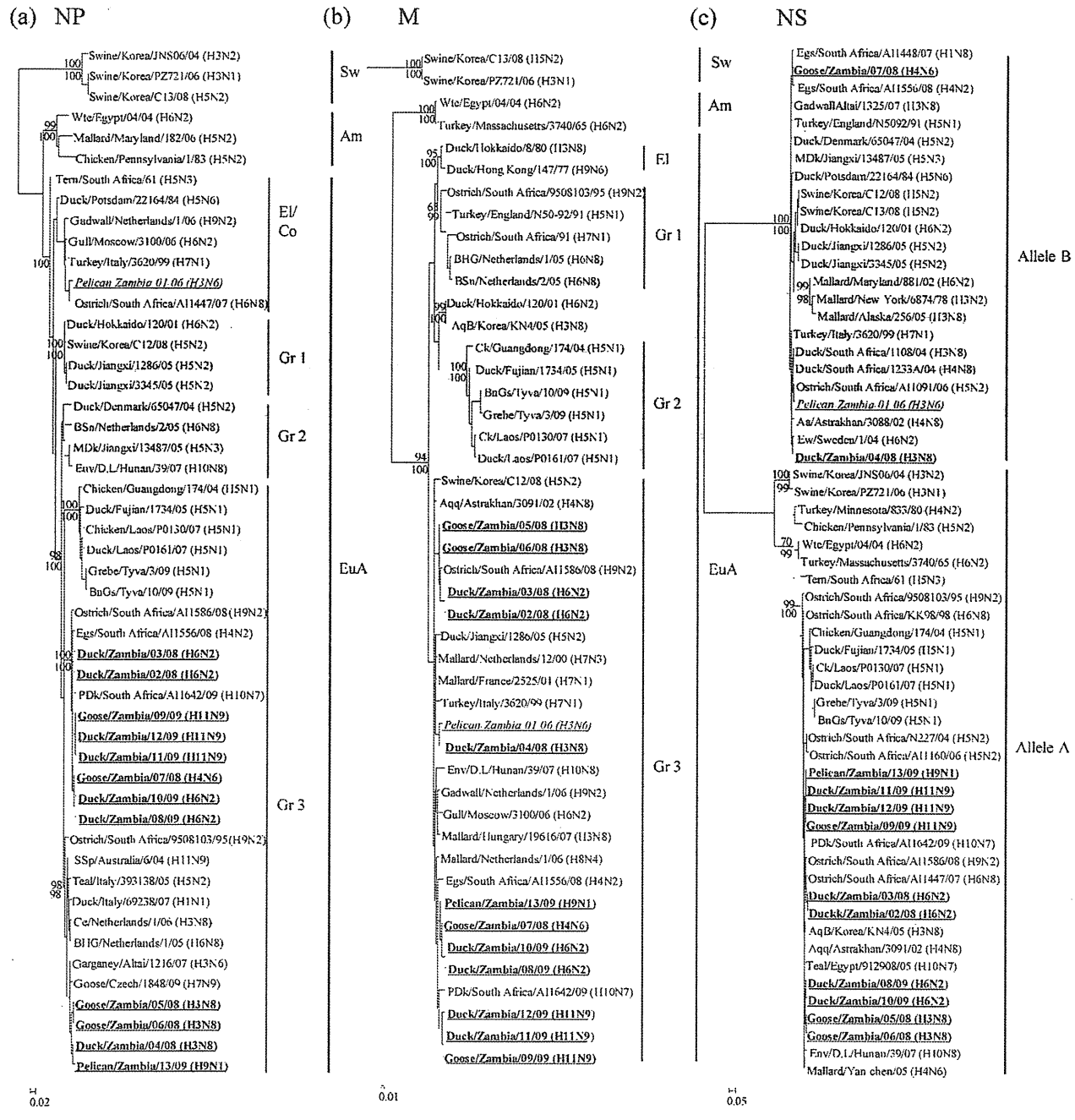

Fig. 4. Phylogenetic relationships of the NP (a), M (b) and NS (c) genes of AIVs isolated from wild birds in Zambia. Analysis was based on nt 46–1489 (1444 bp) of NP, 32–753 (722 bp) of M and 57–705 (649 bp) of NS. Numbers above and below the branch nodes indicate neighbour-joining bootstrap values of ⩾50 % and Bayesian posterior probabilities of >95, respectively. Owing to space limitations, not all supports are indicated. Virus strains characterized in the present study are in bold and underlined, whilst the first influenza virus isolate from an avian host in Zambia is italicized. Bars, number of substitutions per site. Lineages: El/Co, early/contemporary. Strain names: Aa, *Anas angustirostris*; AqB, aquatic bird; Aqq, *Anas querquedula*. Other abbreviations are listed in the legends of Figs 1–3.

Phylogenetic analysis of the non-structural (NS) gene indicated that the NS genes of ten of the viruses from wild birds in Zambia comprised the A allele, whilst the other three were of the B allele (Fig. 4c). Some of the NS genes were closely related to viruses isolated mainly in Asia and Africa, particularly those isolated in South Africa (Fig. 4c).

The NS gene tree clearly demonstrated that, among the viruses examined, the A allele was predominant and that two genetically distinct gene pools, corresponding to NS alleles A and B, were co-circulating in wild birds in this region during the surveillance period.

Amino acid sequence analysis

Although it is difficult to ascertain the capacity of a non-pathogenic AIV/LPAIV from wild waterfowl to cause interspecies transmission into other animals, close monitoring of host-associated signatures in viral proteins may provide some clues regarding an isolate's zoonotic potential. Several amino acids that are preferentially associated with human influenza viruses have been described (Chen *et al.*, 2006; Finkelstein *et al.*, 2007; Shaw *et al.*, 2002). We examined the deduced amino acid sequences of all the internal proteins of all the wild-bird isolates from Zambia and identified some human-associated amino acids in the genome of some strains (Table 2). Zb08 (H6N2) and Zb10 (H6N2) possessed the human-associated amino acid methionine at position 475 of the PB2 protein, which has been described to be 100 % conserved in the influenza viruses that caused the 1918, 1957 and 1968 human pandemics (Finkelstein *et al.*, 2007). Zb13 (H9N1) had a serine at position 66 of the PB1-F2 polypeptide, which was shown previously to contribute to increased virulence in mice (Conenello *et al.*, 2007). Zb13 (H9N1) also possessed the human-associated amino acid alanine at position 76 of the PB1-F2 protein. Six isolates were found to have the human-associated amino acid serine at position 82 of the PB1-F2 protein, whilst only Zb08 (H6N2) and Zb10 (H6N2) had the human-associated amino acid glycine at position 87 of this polypeptide. In the M2 protein, Zb04 (H3N8) possessed the human-associated amino acid valine at position 28. At position 55 of the M2 protein, Zb07 (H4N6), Zb08 (H6N2), Zb10 (H6N2) and Zb12 (H11N9) were found to possess the human-associated amino acid phenylalanine. It is noteworthy that, although the human-associated amino acids found in some of the virus isolates analysed in this report are not unique to these isolates, these residues are rarely found among AIVs isolated from members of the orders Anseriformes and Charadriiformes (our unpublished data).

Table 2. Human-associated amino acids identified in viral proteins of AIVs isolated in Zambia

Protein	Aa position*	Host		Isolate†
		Avian	Human	
PB2	475	L	M	Zb08 (H6N2)
				Zb10 (H6N2)
PB1-F2	66	N	S‡	Zb13 (H9N1)
	76	V	A	Zb13 (H9N1)
	82	L	S	Zb04 (H3N8)
				Zb05 (H3N8)
				Zb06 (H3N8)
				Zb08 (H6N2)
				Zb10 (H6N2)
				Zb13 (H9N1)
	87	E	G	Zb08 (H6N2)
				Zb10 (H6N2)
M2	28	I	V	Zb04 (H3N8)
	55	L	F	Zb07 (H4N6)
				Zb08 (H6N2)
				Zb10 (H6N2)
				Zb12 (H11N9)

*For references of human-associated residues at these specific positions, see Chen *et al.* (2006), Finkelstein *et al.* (2007) and Shaw *et al.* (2002).

†Names of isolates possessing human-associated amino acid residues.

‡The amino acid serine at position 66 of the PB1-F2 protein is not a human-associated residue but was shown previously to increase virulence in mice (Conenello *et al.*, 2007).

Replication and pathogenicity of selected viruses in mice

Amino acid sequence analysis revealed that several isolates from wild birds in Zambia had human-associated residues in their genome (Table 2). Therefore, we sought to investigate whether there could be a difference in virus replication and/or pathogenicity in a mammalian host between viruses either possessing or lacking human-associated residues. For this purpose, we compared the replication ability and pathogenicity of two isolates, Zb03 (H6N2) and Zb10 (H6N2), in mice. Four human-associated residues were identified in some viral proteins of Zb10 (H6N2), whilst none was observed in the genome of Zb03 (H6N2) (Table 2). We also tested the replication capacity and pathogenicity of Zb04 (H3N8) in mice, because it had two human-associated residues in its genome and was of a subtype distinct from that of Zb10 (H6N2).

All the tested viruses replicated in the lungs of mice without prior adaptation, with virus titres ranging from $10^{3.3}$ to $10^{4.8}$ EID_{50} g^{-1} (Table 3). None of the viruses was detected in the brain. It was noted that mice inoculated with Zb10 (H6N2) showed higher virus titres that were statistically significantly different from those of Zb03 (H6N2)-infected mice (Table 3). Virus was detected in the lungs of all five mice inoculated with Zb04 (H3N8) and Zb10 (H6N2), whilst, in Zb03 (H6N2)-inoculated mice, virus was detected in three of the five mice.

Mice infected with Zb10 (H6N2) exhibited more weight loss and delayed weight gain [weight returned to baseline after day 7 post-inoculation (p.i.)] than those inoculated with Zb03 (H6N2) (Fig. 5). Zb04 (H3N8)-inoculated mice showed significant weight loss early on p.i. when compared with Zb03 (H6N2)- or mock-inoculated control mice (Fig. 5). Mild to considerable ruffled fur was noted between days 1 and 3 p.i. in mice infected with Zb10 (H6N2) and Zb04 (H3N8) but not in Zb03 (H6N2)-inoculated mice.

Table 3. Replication of selected AIVs isolated from wild waterfowl in Zambia in BALB/c mice

Virus	No. positive/ total	Mean virus titre of positive samples (log_{10} EID_{50} g^{-1})		*P* value
		Lung	Brain	
Zb03 (H6N2)*	3/5	3.3	$<10^{1.5}$	–
Zb04 (H3N8)†	5/5	3.6	$<10^{1.5}$	0.099
Zb10 (H6N2)†	5/5	4.8	$<10^{1.5}$	0.001‡

*Virus with no apparent human/mammalian-associated residues in its genome.
†Viruses with human-associated residues in their genome.
‡Virus titre in the lungs of mice inoculated with Zb10 (H6N2) was significantly higher than that of Zb03 (H6N2)-inoculated mice (Student's *t*-test, $P<0.05$).

All the mice survived the infection for the 14-day observation period.

DISCUSSION

In this study, we genetically and biologically characterized AIVs isolated from wild birds in Zambia. During the surveillance period, AIVs were isolated mainly between June and November, a time frame encompassing the period when palearctic migrants are absent or rare, as well as when they are present. Palearctic birds usually start to arrive in Zambia between September and December and leave between January and May. Our isolation of AIVs between June and August of 2008 and 2009 when palearctic migrants were scarce raises the possibility of yearly persistence of AIVs in indigenous waterfowl in southern Africa. This idea is further supported by our phylogenetic analyses, which showed the separate clustering of southern African isolates, with the glycoprotein genes of H11N9 viruses characterized in this report forming a distinct sublineage within the Eurasian lineage (Fig. 3a, b and Supplementary Figs S1c and S2c). In neighbouring Zimbabwe, AIVs were also detected in Afro-tropical waterfowl in periods when palearctic birds were rare (Caron *et al.*, 2010). Moreover, AIVs were detected from Afro-tropical bird species in several major wetlands in Africa (Gaidet *et al.*, 2007). These data not only support the notion of a possible endemicity of AIVs in Afro-tropical ecosystems where high temperatures experienced in these regions may restrict the persistence and transmissibility of AIVs (Brown *et al.*, 2009), but also raise the possibility that palearctic migrants may also carry AIVs from Africa into Eurasia. However, the extent to which Afro-tropical ecosystems depend on introductions of AIVs by Eurasian migrants to sustain the possible endemic state remains to be clarified.

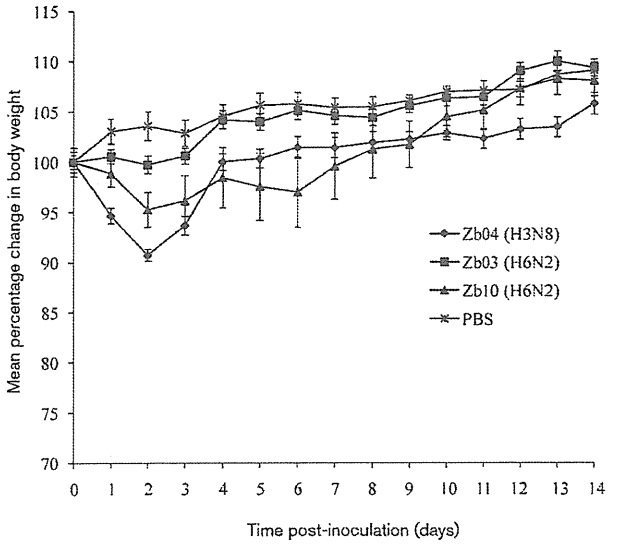

Fig. 5. Weight loss in mice inoculated with selected AIVs from wild waterfowl in Zambia. Data are presented as mean body weight change per group ±SD. Statistically significant weight loss (Student's *t*-test, $P<0.05$) was observed in Zb04 (H3N8)-inoculated mice at days 1 ($P=0.03$), 2 ($P=0.01$) and 3 ($P=0.03$) p.i. compared with mock-inoculated control mice.

The detection of five distinct HA and NA subtypes suggested that a variety of subtypes could be circulating in wild birds in this region. Whilst 11 of the isolates were detected in wild ducks and geese, confirming the major role of these birds in the perpetuation of AIVs (Olsen *et al.*, 2006; Webster *et al.*, 1992), Zb13 (H9N1) was isolated from an atypical avian host, a great white pelican. Despite several AIV surveillance studies that involved sampling from the Pelecaniformes worldwide (Gaidet *et al.*, 2007; Munster *et al.*, 2007; Olsen *et al.*, 2006), the number of AIVs detected from this order has remained low. Thus, we consider the two instances in which we isolated AIVs from these birds as incidental findings, but we do not exclude the possibility that white pelicans, which are native to southern Africa, may also play a major role in influenza virus ecology in this region.

Phylogenetic analyses demonstrated that all the gene segments of the viruses reported in this study clustered with contemporary viruses of the Eurasian avian lineage. Most genes were closely related to those of AIVs isolated from wild and domestic birds in South Africa. AIVs originating in wild birds have been implicated in avian influenza outbreaks in farmed birds in South Africa with serious economic consequences (Abolnik, 2007; Abolnik *et al.*, 2007, 2010; Alexander, 2007; Brown, 2010). These data highlight the need for continued monitoring of AIVs in wild and domestic birds in southern Africa for avian influenza control. It is also important to clarify the extent of influenza virus exchange between wild birds and domesticated birds (including ostriches) in the region, as

results from a study in China demonstrated that a two-way transmission of influenza viruses between terrestrial and aquatic birds may increase opportunities for the generation of reassortant viruses with pandemic potential (Li *et al.*, 2003). Furthermore, the potential role of human related activities (e.g. the poultry trade) in AIV dissemination should not be ignored.

A number of human-associated amino acids were observed in some viral proteins of some viruses tested. The two possible means by which AIVs may acquire 'novel' amino acids are either through genetic reassortment or through point mutations. Genetic analyses of AIVs isolated from wild and terrestrial birds in southern Africa have demonstrated the involvement of ostriches in the evolution and epidemiology of AIVs in this region (Abolnik, 2007; Abolnik *et al.*, 2007, 2010; this study). Recently, Shinya *et al.* (2009) demonstrated that ostriches may be involved in the emergence of viruses possessing mammalian-associated amino acids lysine and asparagine at positions 627 and 701 of the PB2 protein, respectively. Indeed, an examination of PB2 gene sequences of viruses isolated from ostriches in South Africa between 1995 and 2008 showed that four viruses had lysine and one virus possessed asparagine at positions 627 and 701 of the PB2 protein, respectively (data not shown). Therefore, if a two-way transmission of AIVs between ostriches and wild aquatic birds in southern Africa exists, these data indicate that the human-associated amino acids observed in some internal proteins of some of the isolates examined here may have been acquired through genetic reassortment with viruses from ostriches. Unfortunately, the lack of complete internal protein gene sequences of isolates from ostriches in South Africa for a comprehensive study makes it difficult to reach this conclusion. Moreover, genetic analyses of the deduced amino acids revealed that the surface proteins of the viruses listed in Table 2 maintained typical features of non-pathogenic wild waterfowl isolates, including conservation of putative glycosylation sites and no NA stalk deletions, and did not exhibit evidence for accelerated or increased amino acid substitutions, suggesting that these viruses may not have circulated extensively in land-based avian species. These observations leave open the possibility that the human-associated amino acids in the viral proteins of some isolates from Zambia may have been acquired in wild waterfowl or other non-gallinaceous birds. Whether some African waterfowl may provide an environment that may lead to the selection of AIVs with human/mammalian-associated amino acids is a question deserving further exploration.

In a mouse model, we demonstrated that all the tested viruses replicated in mouse lung without prior adaptation and that mice infected with isolates having human-associated residues displayed increased virus titres and caused increased morbidity, as measured by weight loss, than those inoculated with Zb03 (H6N2). Although it is tempting to conclude that possession of human-associated residues may have impacted on virus replication and pathogenicity in mice, there is need for caution, because the influence of other residues was not ruled out in the current study. In fact, there were 68 amino acid differences in viral proteins between Zb03 (H6N2) and Zb10 (H6N2). Therefore, investigations employing reverse genetics and site-directed mutagenesis may be needed to explain more fully the observed differences. To our knowledge, the present study is the first to demonstrate the ability of non-HPAIVs from wild birds in Africa to replicate without adaptation and cause illness in a mammalian host. Elsewhere, although few in number, AIVs from wild birds of considerable numbers of HA subtypes have been shown to replicate in mice and ferrets without adaptation, causing varied degrees of morbidity (Driskell *et al.*, 2010; Gillim-Ross *et al.*, 2008; Joseph *et al.*, 2007; Kim *et al.*, 2010; Wan *et al.*, 2008). These studies have highlighted the potential risk of direct transmission of non-HPAIVs from wild birds to mammalian species. Whilst direct transmission of AIVs from wild birds to humans has not been reported, serological evidence of AIV infection in three persons with substantial exposure to wild waterfowl and game birds argues for a possible direct transmission of AIVs from wild birds to humans (Gill *et al.*, 2006). Moreover, both natural and experimental infections of humans with AIVs, together with serological data, have emphasized the susceptibility of humans to several AIV subtypes (Myers *et al.*, 2007; Peiris *et al.*, 2007; Shortridge, 1992). Thus, the potential threat posed to both animal and public health by some of the viruses characterized currently cannot be overemphasized.

Here, we demonstrated that the 12 influenza viruses isolated from wild waterfowl in Zambia belonged to the contemporary Eurasian avian lineage. We have shown the possibility that AIVs could persist in wild waterfowl in a Zambian ecosystem, with transmission of viruses involving wild and domestic avian species in southern Africa, Europe and Asia. This study further established that some AIVs from wild waterfowl in Zambia may have the potential to infect mice directly without adaptation. Overall, the present study raises concerns for continued monitoring of AIVs in wild and domestic birds in southern Africa and suggests that complete characterization of isolates may help in the identification of strains that may have potential for future incursions into humans and other animals.

METHODS

Viruses and sequencing. The viruses characterized in the present study were isolated from wild waterfowl faecal specimens collected in Lochinvar National Park between April 2008 and November 2009 (Table 1). All virus isolation was performed using 10–11-day-old embryonated chicken's eggs. The isolates were subtyped by standard HA inhibition and NA inhibition tests, as well as by sequencing of the HA and NA genes. The viruses were passaged once in eggs before being used in this study. Viral RNA extraction, cDNA synthesis, PCR and sequencing were carried out as described previously (Simulundu *et al.*, 2009).

Phylogenetic analyses. Phylogenetic trees were constructed by the neighbour-joining bootstrap method with 1000 replicates applied

using MEGA4 (Tamura *et al.*, 2007). The gene tree topologies obtained in MEGA4 were then confirmed using Bayesian methods implemented in MRBAYES version 3.1.2 (Huelsenbeck & Ronquist, 2001). Specifically, we used the program ModelTest version 3.7 (Posada & Crandall, 2001), applied in PAUP* version 4.0 (Swofford, 2001), to determine the appropriate evolutionary model that best fitted the data. The HA, NA, PB2, PB1, PA and NP nucleotide sequence data were best fitted by the general time reversible plus invariant sites plus gamma-distributed (GTR+I+G) model, whilst the Hasegawa–Kishino–Yano (plus invariant sites) plus gamma-distributed models (HKY+G and HKY+I+G) were preferred for the NS and M sequence data, respectively. In Bayesian analysis, we used one to four replicates of 1 million generations, with four chains sampled every 100 generations. All replicates converged with less than 0.01 SD of split frequencies.

Experimental infection of mice. Groups of 6-week-old BALB/c mice (ten mice per group) were lightly anaesthetized with isoflurane and inoculated intranasally with 0.05 ml virus-infected chorioallantoic fluid containing Zb03 (H6N2), Zb04 (H3N8) or Zb10 (H6N2) ($10^{7.5}$ EID_{50} ml^{-1}). To serve as a control, a group of five mice was mock infected with sterile PBS. Mice were observed daily for morbidity (weight loss, ruffled fur and hunching) and mortality for 14 days. On day 3 p.i., half of the virus-inoculated mice were euthanized, and the titres of virus in the lung and brain were determined using eggs. Briefly, a 10% lung and brain tissue homogenate was prepared using minimal essential medium (Gibco) containing antibiotics. The tissue homogenates were clarified by centrifugation and titrated in 10–11-day-old embryonated chicken's eggs. The virus titre was calculated as the $\log_{10}$ EID_{50} (g tissue)$^{-1}$ by the method of Reed & Muench (1938).

ACKNOWLEDGEMENTS

We thank the Zambia Wildlife Authority for supporting the wild-bird influenza A virus surveillance programme in Zambia. We are also grateful to H. Miyamoto, A. Ohnuma and A. Yokoyama for excellent technical assistance. This work was supported by the Program of Founding Research Centers for Emerging and Reemerging Infectious Diseases and the Global COE Program 'Establishment of International Collaboration Centers for Zoonosis Control' from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

REFERENCES

Abolnik, C. (2007). Molecular characterization of H5N2 avian influenza viruses isolated from South African ostriches in 2006. *Avian Dis* **51**, 873–879.

Abolnik, C., Bisschop, S., Gerdes, T., Olivier, A. & Horner, R. (2007). Outbreaks of avian influenza H6N2 viruses in chickens arose by a reassortment of H6N8 and H9N2 ostrich viruses. *Virus Genes* **34**, 37–45.

Abolnik, C., Gerdes, G. H., Sinclair, M., Ganzevoort, B. W., Kitching, J. P., Burger, C. E., Romito, M., Dreyer, M., Swanepoel, S. & other authors (2010). Phylogenetic analysis of influenza A viruses (H6N8, H1N8, H4N2, H9N2, H10N7) isolated from wild birds, ducks, and ostriches in South Africa from 2007 to 2009. *Avian Dis* **54** (Suppl.), 313–322.

Alexander, D. J. (2007). Summary of avian influenza activity in Europe, Asia, Africa, and Australasia, 2002-2006. *Avian Dis* **51** (Suppl.), 161–166.

Bahl, J., Vijaykrishna, D., Holmes, E. C., Smith, G. J. D. & Guan, Y. (2009). Gene flow and competitive exclusion of avian influenza A virus in natural reservoir hosts. *Virology* **390**, 289–297.

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

Brown, I. H. (2010). Summary of avian influenza activity in Europe, Asia, and Africa, 2006–2009. *Avian Dis* **54** (Suppl.), 187–193.

Brown, J. D., Goekjian, G., Poulson, R., Valeika, S. & Stallknecht, D. E. (2009). Avian influenza virus in water: infectivity is dependent on pH, salinity and temperature. *Vet Microbiol* **136**, 20–26.

Capua, I. & Alexander, D. J. (2006). The challenge of avian influenza to the veterinary community. *Avian Pathol* **35**, 189–205.

Caron, A., Abolnik, C., Mundava, J., Gaidet, N., Burger, C. E., Mochotlhoane, B., Bruinzeel, L., Chiweshe, N., de Garine-Wichatitsky, M. & Cumming, G. S. (2010). Persistence of low pathogenic avian influenza virus in waterfowl in a Southern African ecosystem. *EcoHealth* (Epub ahead of print).

Chen, G.-W., Chang, S.-C., Mok, C.-K., Lo, Y.-L., Kung, Y.-N., Huang, J.-H., Shih, Y.-H., Wang, J.-Y., Chiang, C. & other authors (2006). Genomic signatures of human versus avian influenza A viruses. *Emerg Infect Dis* **12**, 1353–1360.

Conenello, G. M., Zamarin, D., Perrone, L. A., Tumpey, T. & Palese, P. (2007). A single mutation in the PB1-F2 of H5N1 (HK/97) and 1918 influenza A viruses contributes to increased virulence. *PLoS Pathog* **3**, 1414–1421.

de Wit, E., Kawaoka, Y., de Jong, M. D. & Fouchier, R. A. (2008). Pathogenicity of highly pathogenic avian influenza virus in mammals. *Vaccine* **26** (Suppl. 4), D54–D58.

Driskell, E. A., Jones, C. A., Stallknecht, D. E., Howerth, E. W. & Tompkins, S. M. (2010). Avian influenza virus isolates from wild birds replicate and cause disease in a mouse model of infection. *Virology* **399**, 280–289.

Duan, L., Campitelli, L., Fan, X. H., Leung, Y. H., Vijaykrishna, D., Zhang, J. X., Donatelli, I., Delogu, M., Li, K. S. & other authors (2007). Characterization of low-pathogenic H5 subtype influenza viruses from Eurasia: implications for the origin of highly pathogenic H5N1 viruses. *J Virol* **81**, 7529–7539.

Ducatez, M. F., Olinger, C. M., Owoade, A. A., De Landtsheer, S., Ammerlaan, W., Niesters, H. G. M., Osterhaus, A. D. M. E., Fouchier, R. A. M. & Muller, C. P. (2006). Avian flu: multiple introductions of H5N1 in Nigeria. *Nature* **442**, 37.

Finkelstein, D. B., Mukatira, S., Mehta, P. K., Obenauer, J. C., Su, X., Webster, R. G. & Naeve, C. W. (2007). Persistent host markers in pandemic and H5N1 influenza viruses. *J Virol* **81**, 10292–10299.

Gaidet, N., Dodman, T., Caron, A., Balança, G., Desvaux, S., Goutard, F., Cattoli, G., Lamarque, F., Hagemeijer, W. & Monicat, F. (2007). Avian influenza viruses in water birds, Africa. *Emerg Infect Dis* **13**, 626–629.

Gaidet, N., Cattoli, G., Hammoumi, S., Newman, S. H., Hagemeijer, W., Takekawa, J. Y., Cappelle, J., Dodman, T., Joannis, T. & other authors (2008). Evidence of infection by H5N2 highly pathogenic avian influenza viruses in healthy wild waterfowl. *PLoS Pathog* **4**, e1000127.

Gill, J. S., Webby, R., Gilchrist, M. J. R. & Gray, G. C. (2006). Avian influenza among waterfowl hunters and wildlife professionals. *Emerg Infect Dis* **12**, 1284–1286.

Gillim-Ross, L., Santos, C., Chen, Z., Aspelund, A., Yang, C.-F., Ye, D., Jin, H., Kemble, G. & Subbarao, K. (2008). Avian influenza H6 viruses productively infect and cause illness in mice and ferrets. *J Virol* **82**, 10854–10863.

Hinshaw, V. S., Webster, R. G., Easterday, B. C. & Bean, W. J., Jr (1981). Replication of avian influenza A viruses in mammals. *Infect Immun* **34**, 354–361.

Huelsenbeck, J. P. & Ronquist, F. R. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.

Joseph, T., McAuliffe, J., Lu, B., Jin, H., Kemble, G. & Subbarao, K. (2007). Evaluation of replication and pathogenicity of avian influenza a H7 subtype viruses in a mouse model. *J Virol* **81**, 10558–10566.

Kida, H., Ito, T., Yasuda, J., Shimizu, Y., Itakura, C., Shortridge, K. F., Kawaoka, Y. & Webster, R. G. (1994). Potential for transmission of avian influenza viruses to pigs. *J Gen Virol* **75**, 2183–2188.

Kim, H.-R., Lee, Y.-J., Lee, K.-K., Oem, J.-K., Kim, S.-H., Lee, M. -H., Lee, O.-S. & Park, C.-K. (2010). Genetic relatedness of H6 subtype avian influenza viruses isolated from wild birds and domestic ducks in Korea and their pathogenicity in animals. *J Gen Virol* **91**, 208–219.

Li, K. S., Xu, K. M., Peiris, J. S., Poon, L. L., Yu, K. Z., Yuen, K. Y., Shortridge, K. F., Webster, R. G. & Guan, Y. (2003). Characterization of H9 subtype influenza viruses from the ducks of southern China: a candidate for the next influenza pandemic in humans? *J Virol* **77**, 6988–6994.

Li, K. S., Guan, Y., Wang, J., Smith, G. J. D., Xu, K. M., Duan, L., Rahardjo, A. P., Puthavathana, P., Buranathai, C. & other authors (2004). Genesis of a highly pathogenic and potentially pandemic H5N1 influenza virus in eastern Asia. *Nature* **430**, 209–213.

Munster, V. J., Baas, C., Lexmond, P., Waldenström, J., Wallensten, A., Fransson, T., Rimmelzwaan, G. F., Beyer, W. E., Schutten, M. & other authors (2007). Spatial, temporal, and species variation in prevalence of influenza A viruses in wild migratory birds. *PLoS Pathog* **3**, e61.

Murphy, B. R., Hinshaw, V. S., Sly, D. L., London, W. T., Hosier, N. T., Wood, F. T., Webster, R. G. & Chanock, R. M. (1982). Virulence of avian influenza A viruses for squirrel monkeys. *Infect Immun* **37**, 1119–1126.

Myers, K. P., Setterquist, S. F., Capuano, A. W. & Gray, G. C. (2007). Infection due to 3 avian influenza subtypes in United States veterinarians. *Clin Infect Dis* **45**, 4–9.

Olsen, B., Munster, V. J., Wallensten, A., Waldenström, J., Osterhaus, A. D. & Fouchier, R. A. (2006). Global patterns of influenza A virus in wild birds. *Science* **312**, 384–388.

Peiris, J. S., de Jong, M. D. & Guan, Y. (2007). Avian influenza virus (H5N1): a threat to human health. *Clin Microbiol Rev* **20**, 243–267.

Posada, D. & Crandall, K. A. (2001). Selecting the best-fit model of nucleotide substitution. *Syst Biol* **50**, 580–601.

Reed, L. J. & Muench, H. (1938). A simple method of estimating fifty percent endpoints. *Am J Hyg* **27**, 493–497.

Röhm, C., Horimoto, T., Kawaoka, Y., Süss, J. & Webster, R. G. (1995). Do hemagglutinin genes of highly pathogenic avian influenza viruses constitute unique phylogenetic lineages? *Virology* **209**, 664–670.

Shaw, M., Cooper, L., Xu, X., Thompson, W., Krauss, S., Guan, Y., Zhou, N., Klimov, A., Cox, N. & other authors (2002). Molecular changes associated with the transmission of avian influenza a H5N1 and H9N2 viruses to humans. *J Med Virol* **66**, 107–114.

Shinya, K., Makino, A., Ozawa, M., Kim, J. H., Sakai-Tagawa, Y., Ito, M., Le, Q. M. & Kawaoka, Y. (2009). Ostrich involvement in the selection of H5N1 influenza virus possessing mammalian-type amino acids in the PB2 protein. *J Virol* **83**, 13015–13018.

Shortridge, K. F. (1992). Pandemic influenza: a zoonosis? *Semin Respir Infect* **7**, 11–25.

Simulundu, E., Mweene, A. S., Tomabechi, D., Hang'ombe, B. M., Ishii, A., Suzuki, Y., Nakamura, I., Sawa, H., Sugimoto, C. & other authors (2009). Characterization of H3N6 avian influenza virus isolated from a wild white pelican in Zambia. *Arch Virol* **154**, 1517–1522.

Smith, G. J. D., Fan, X. H., Wang, J., Li, K. S., Qin, K., Zhang, J. X., Vijaykrishna, D., Cheung, C. L., Huang, K. & other authors (2006). Emergence and predominance of an H5N1 influenza variant in China. *Proc Natl Acad Sci U S A* **103**, 16936–16941.

Swofford, D. L. (2001). PAUP*: phylogenetic analysis using parsimony (and other methods) 4.0 beta. Sunderland, MA: Sinauer Associates.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596–1599.

Wan, H., Sorrell, E. M., Song, H., Hossain, M. J., Ramirez-Nieto, G., Monne, I., Stevens, J., Cattoli, G., Capua, I. & other authors (2008). Replication and transmission of H9N2 influenza viruses in ferrets: evaluation of pandemic potential. *PLoS ONE* **3**, e2923.

Wang, G., Zhan, D., Li, L., Lei, F., Liu, B., Liu, D., Xiao, H., Feng, Y., Li, J. & other authors (2008). H5N1 avian influenza re-emergence of Lake Qinghai: phylogenetic and antigenic analyses of the newly isolated viruses and roles of migratory birds in virus circulation. *J Gen Virol* **89**, 697–702.

Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M. & Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. *Microbiol Rev* **56**, 152–179.

World Health Organization (2010). Cumulative number of confirmed human cases of avian influenza A/(H5N1) reported to WHO. http://www.who.int/csr/disease/avian_influenza/country/cases_table_2010_12_29/en/index.html. Accessed 4 January 2011.

Xu, K. M., Smith, G. J. D., Bahl, J., Duan, L., Tai, H., Vijaykrishna, D., Wang, J., Zhang, J. X., Li, K. S. & other authors (2007). The genesis and evolution of H9N2 influenza viruses in poultry from southern China, 2000 to 2005. *J Virol* **81**, 10389–10401.

Antibody-Dependent Enhancement of Marburg Virus Infection

Eri Nakayama,[1] Daisuke Tomabechi,[1] Keita Matsuno,[1] Noriko Kishida,[2] Reiko Yoshida,[1] Heinz Feldmann,[3] and Ayato Takada[1]

[1]Department of Global Epidemiology, Hokkaido University Research Center for Zoonosis Control, Sapporo, [2]Laboratory of Influenza Virus Surveillance, Center for Influenza Virus Research, National Institute of Infectious Diseases, Tokyo, Japan; and [3]Laboratory of Virology, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rocky Mountain Laboratories, Hamilton, Montana

Background. Marburg virus (MARV) and Ebola virus (EBOV) cause severe hemorrhagic fever in primates. Earlier studies demonstrated that antibodies to particular epitopes on the glycoprotein (GP) of EBOV enhanced virus infectivity in vitro.

Methods. To investigate this antibody-dependent enhancement (ADE) in MARV infection, we produced mouse antisera and monoclonal antibodies (mAbs) to the GPs of MARV strains Angola and Musoke.

Results. The infectivity of vesicular stomatitis virus pseudotyped with Angola GP in K562 cells was significantly enhanced in the presence of Angola GP antisera, whereas only minimal ADE activity was seen with Musoke GP antisera. This difference correlated with the percentage of hybridoma clones producing infectivity-enhancing mAbs. Using mAbs to MARV GP, we identified 3 distinct ADE epitopes in the mucinlike region on Angola GP. Interestingly, some of these antibodies bound to both Angola and Musoke GPs but showed significantly higher ADE activity for strain Angola. ADE activity depended on epitopes in the mucinlike region and glycine at amino acid position 547, present in the Angola but absent in the Musoke GP.

Conclusions. These results suggest a possible link between ADE and MARV pathogenicity and provide new insights into the mechanisms underlying ADE entry of filoviruses.

Marburg virus (MARV) and Ebola virus (EBOV) are filamentous, enveloped, negative-strand RNA viruses belonging to the family Filoviridae. These viruses have produced sporadic outbreaks of hemorrhagic fever in human and nonhuman primates [1]. MARV was first identified in 1967 during an outbreak of hemorrhagic fever in Marburg, Germany, and Belgrade, Yugoslavia. These outbreaks were linked to infected monkeys imported from Uganda [2]. Since the first outbreak of Marburg hemorrhagic fever, several sporadic outbreaks have been reported in central African countries [1, 3–6]. The largest outbreak occurred in Uige province in Angola from 2004 to 2005, with a mortality rate of 90% among 252 reported cases [7]. The strain Angola had a higher mortality rate and was thought to be more pathogenic than earlier isolates, such as strain Musoke [8, 9], which was isolated from a human case in 1980 in Kenya [5]. Among EBOV species, *Zaire ebolavirus* seems to be the most virulent, with a case fatality rate of up to 90%; whereas *Reston ebolavirus* has never been associated with symptomatic infection in humans [10] and was shown to be less pathogenic than *Zaire ebolavirus* in nonhuman primates [11].

The filovirus genome encodes at least 7 structural proteins. The fourth gene from the 3' end of the genome encodes the envelope glycoprotein (GP), which undergoes proteolytic cleavage into 2 subunits, GP1 and GP2. The GP1 subunit mediates cell-surface receptor binding [12, 13], and the GP2 subunit is involved in fusion of the viral envelope and host cell membrane [14]. GP is highly glycosylated, with a large amount of N- and O-linked

Potential conflicts of interest: none reported.

Presented in part: 57th Annual Meeting of the Japanese Society for Virology, Tokyo, Japan, 25–27 October 2009.

Correspondence: Ayato Takada, PhD, Department of Global Epidemiology, Hokkaido University Research Center for Zoonosis Control, Kita-20, Nishi-10, Kita-ku, Sapporo 001-0020, Japan (atakada@czc.hokudai.ac.jp).

The Journal of Infectious Diseases 2011;204:S978–S985

0022-1899 (print)/1537-6613 (online)/2011/204S3-0033$14.00
DOI: 10.1093/infdis/jir334

glycans, most of which are located in the middle portion of the GP, designated the mucinlike region (MLR) [15, 16]. The amino acid sequences of the MLR are highly variable among filovirus species [17, 18].

It has been demonstrated elsewhere that EBOV infection in humans and nonhuman primates induced GP-specific antibodies that had the ability to enhance viral infectivity of certain cells in vitro [19, 20]. This mechanism, known as antibody-dependent enhancement (ADE) of viral infection, depends mostly on the cross-linking of virus-antibody complexes through interaction with cellular Fc receptors (FcRs) [21]. Our previous studies have identified an additional mechanism underlying EBOV ADE in vitro, namely, complement protein C1q and C1q receptor-dependent ADE [19]. Epitopes involved in ADE were identified predominantly in the MLR of the Zaire EBOV (ZEBOV) GP1 subunit. A possible contribution of ADE to the distinct pathogenicity observed for ZEBOV and Reston EBOV was discussed elsewhere [20, 22]. However, little is known about the role of ADE in MARV pathogenicity.

In this study, we produced mouse antisera and monoclonal antibodies (mAbs) specific to the GPs of MARV strains Angola and Musoke and examined their ADE activities by using vesicular stomatitis virus (VSV) pseudotyped with MARV GPs. We found distinct ADE activity between antibodies to Angola and Musoke GPs, which may be linked to a difference in pathogenicity of these strains. The identified ADE epitopes were all located in the MLR of the GPs, but the presence of these epitopes was not sufficient to give a maximal ADE. Potential mechanisms for effective ADE seen with certain MARV strains are discussed here.

MATERIALS AND METHODS

Viruses and Cells

VSV pseudotyped with Angola GP (VSV-Angola) or Musoke GP (VSV-Musoke) expressing green fluorescent protein was generated as described elsewhere [23]. Deletion mutant GPs, chimeric GPs, and mutant GPs with a single substitution were generated as described elsewhere [24]. There was no significant difference in the infectivity in Vero E6 cells among these viruses, suggesting that the functional GPs were incorporated into VSV virions [24]. To reduce the background infectivity of parent VSV G, the pseudotyped viruses were treated with a neutralizing mAb to VSV G protein (VSV-G[N]1-9) before use. The virus infectivity was determined by counting the number of cells expressing green fluorescent protein, using fluorescence microscopy or flow cytometry. Monkey kidney Vero E6 cells and human embryonic kidney (HEK) 293 and 293T cells were grown in Dulbecco's modified Eagle's medium (Sigma), and human chronic myelogenous leukemia K562 cells bearing FcR were grown in Roswell Park Memorial Institute 1640 medium (Sigma). The media were supplemented with fetal calf serum and antibiotics.

Antisera

To produce antisera to filovirus GPs, 5-week-old female BALB/c mice were immunized subcutaneously twice in a 3-week interval with 100 μg of viruslike particles (VLPs) [25, 26] with complete Freund's adjuvant or intraperitoneally twice in a 3-week interval with 50 μg of VLPs only. The serum samples from intraperitoneally immunized mice were collected 7 days after the second immunization. Subcutaneously immunized mice were boosted intraperitoneally with 100 μg of VLPs alone 3 weeks after the second immunization, and the serum samples were collected 7 days after the boost dose.

Generation of mAbs

Five-week-old female BALB/c mice were immunized subcutaneously with 100 μg of VLPs with complete Freund's adjuvant (Difco). At 3 and 6 weeks after the first immunization, the mice were subcutaneously immunized with 100 μg of VLPs with incomplete Freund's adjuvant (Difco). Three weeks after the last immunization, mice were boosted intraperitoneally twice in a 3-week interval with 100 μg of VLPs only. Three days later, mouse spleen cells and mouse myeloma P3-U1 cells were fused and maintained according to a standard procedure [27]. Hybridomas were screened for secretion of MARV GP-specific mAbs by enzyme-linked immunosorbent assay (ELISA), and hybridoma-producing mAbs were cloned by limiting dilution of the cells. mAbs were purified from mouse ascites using protein A agarose columns (Bio-Rad). The isotypes of the obtained mAbs were determined using a mouse mAb isotyping test kit (AbD Serotec) according to the manufacturer's instructions.

Infectivity Enhancement and Neutralization Tests

Appropriately diluted serum samples or mAbs were mixed with equal volumes of the pseudotyped viruses ($\sim 10^5$ infectious units/mL on Vero E6 cells), followed by 1-hour incubation. Infectivity was then determined in Vero E6, K562, and HEK 293 cells for neutralizing, FcR-, and C1q-dependent ADE activities, respectively, by counting the fluorescent cells, as described elsewhere [19, 20, 22, 23]. The relative percentage of infected cells was determined by setting the number of cells infected in the absence of GP-specific antisera or purified mAbs to 100. Antibodies that gave relative infectivity values of $<50\%$ or $>200\%$ were defined as neutralizing or enhancing antibodies, respectively. For detection of C1q-dependent ADE, viruses were incubated with antisera or mAbs in the presence of C1q (50 μg/mL; Sigma) before infection of HEK 293 cells.

Enzyme-Linked Immunosorbent Assay

ELISA plates (Nunc Maxisorp) were coated with lysate from HEK 293T cells expressing MARV-GP, VLP, or histidine-tagged purified GP, as described elsewhere [28], at 4°C overnight and then washed with phosphate-buffered saline (PBS) containing 0.05% Tween 20 (PBST) before addition of blocking buffer (3% skim milk in PBST) for 2 hours at room temperature. After

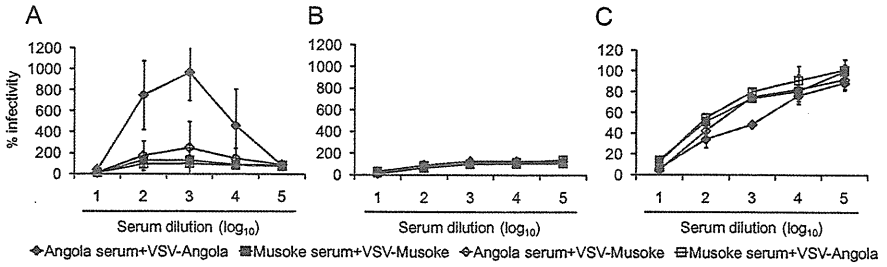

Figure 1. Antibody-dependent enhancement and neutralizing activities of polyclonal antisera from subcutaneously immunized mice. Marburg virus glycoprotein (MARV GP) antisera ($1:10^1$ to $1:10^5$ dilutions) were mixed with vesicular stomatitis virus (VSV) pseudotyped with MARV GPs, incubated for 1 hour at room temperature, and inoculated into 10^5 K562 cells (in 96-well plate) *A*, confluent HEK 293 cells *B*, or confluent Vero E6 cells *C*, Human embryonic kidney 293 cells were infected in the presence of purified C1q (50 μg/mL). Results are expressed as means (± standard deviations) of data for 3 immunized mice. Relative percentages of infected cells are shown as mean values determined based on the number of infected cells in the absence of specific antibodies against MARV GPs (100%).

being washed 3 times with PBST, primary antibodies (ie, hybridoma supernatants, purified mAbs, or antiserum) were added to each well and incubated at room temperature for 1 hour and the plates were washed 3 times with PBST. The antibody binding was detected by goat anti-mouse immunoglobulin (Ig) G1, IgG2a, IgG2b, IgG3 (Bethyl), and IgG (H+L) (Jackson ImmunoResearch) conjugated with horseradish peroxidase. After incubation at room temperature for 1 hour, the plates were washed 4 times with PBST, and 3,3',5,5'-tetramethylbenzidine (Sigma) was added to each well. An equal volume of 1N sulfuric acid was subsequently used to stop the enzyme reaction after 15-minutes incubation, and the optical density value was read at 450 nm on an ELISA plate reader.

Immunostaining

HEK 293T cells were transfected with the mammalian expression plasmid pCAGGS, expressing MARV GPs, and were fixed 48 hours later with methanol for 30 minutes. After blocking with 10% bovine serum albumin in PBS for 90 minutes at room temperature, cells were incubated with mAbs (1 μg/mL) for 1 hour at room temperature. The binding of the mAbs was detected by goat anti-mouse IgG conjugated with horseradish peroxidase (Jackson ImmunoResearch) diluted in 5% bovine serum albumin in PBST. After incubation for 1 hour at room temperature, GP expressed in the cells was visualized with 3,3'-diaminobenzidine.

RESULTS

FcR-Dependent ADE Activity of GP Antisera Differed Between MARV Strains Angola and Musoke

To examine ADE of MARV infection, we first generated antiserum by subcutaneous immunization of mice with MARV strain Angola or Musoke VLPs. The levels of infectivity of VSV-Angola and VSV-Musoke in the presence of these antisera were determined by using K562 and HEK 293 cells for FcR- and C1q-dependent ADE, respectively (Figure 1*A* and 1*B*). In K562 cells, the infectivity of VSV-Angola was markedly enhanced in the presence of Angola GP antiserum at dilutions ranging from $1:10^2$ to $1:10^4$, whereas only minimal enhancement of VSV-Musoke was seen in the presence of Angola GP antiserum. In the presence of Musoke GP antiserum, no apparent enhancement of infectivity was observed for any of the 2 VSV pseudotypes. Furthermore, no C1q-dependent ADE activity in HEK 293 cells was detected in the Angola or Musoke GP antiserum at any dilution (Figure 1*B*), and no significant difference in neutralizing activity was observed between the 2 antisera (Figure 1*C*). Interestingly, cross-reactivity between Angola and Musoke GP antiserum was found in their neutralizing activities but not in FcR-dependent ADE activities.

Difference in C1q-Dependent ADE Activity Between MARV and EBOV Antisera

It was demonstrated elsewhere that intraperitoneal immunization elicited C1q-dependent ADE antibodies to ZEBOV GP more efficiently than subcutaneous immunization [20]. Therefore, we produced MARV strain Angola and Musoke GP antisera by intraperitoneal immunization and compared C1q-dependent ADE activity with ZEBOV GP antisera in HEK 293 cells (Figure 2*A*). We confirmed that ZEBOV GP antisera significantly enhanced the infectivity of VSV pseudotyped with ZEBOV GP in the presence of purified C1q. However, the Angola GP antiserum only minimally enhanced the infectivity of VSV-Angola in the presence of C1q. The Musoke GP antiserum did not show any C1q-dependent ADE activity. Subsequently, we used ELISA to examine the antiserum for the proportion of each IgG subclass (Figure 2*B*). We found that the amount of IgG2a, IgG2b, and IgG3 was significantly lower in the MARV-Angola GP antiserum than in the ZEBOV GP antisera (Student's *t* test, $P < .05$). There was no significant difference in the amount of IgG1 between the MARV-Angola GP and ZEBOV GP antisera.

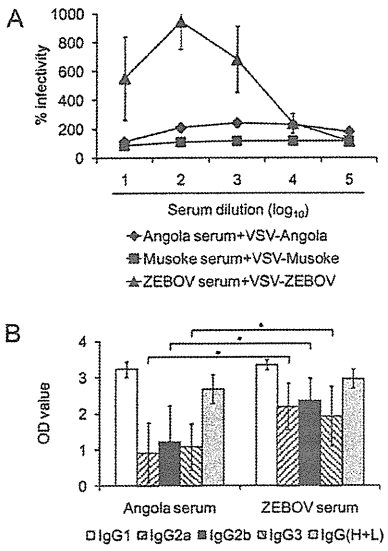

Figure 2. C1q-dependent antibody-dependent enhancement of polyclonal antisera from intraperitoneally immunized mice *A*, and immunoglobulin (Ig) G subclass in the serum samples *B*, Vesicular stomatitis virus (VSV) pseudotyped with Marburg virus glycoproteins (MARV GPs) were incubated with mouse antisera ($1:10^1$ to $1:10^5$ dilutions) in the presence of purified C1q (50 μg/mL) and inoculated into human embryonic kidney 293 cells. Results are expressed as means (± standard deviations) of data for 3 immunized mice. Relative percentages of infected cells are shown as mean values determined based on the number of infected cells in the absence of specific antibodies against MARV GPs (100%). The binding activity of each antibody class in 3 immunized mouse antisera ($1:10^2$) was measured by enzyme-linked immunosorbent assay (ELISA) using histidine-tagged GPs as antigens. Experiments were triplicated and the means and standard deviations were shown. The differences between optical density (OD) values were compared using the Student's *t* test. $^*P < .05$ for differences in OD values between *Angola* and *Zaire ebolavirus* (ZEBOV) GPs.

Characterization of mAbs Specific to MARV GPs

We then generated hybridoma cells producing specific mAbs (47 and 28 clones for Angola and Musoke GPs, respectively). Table 1 lists the number of clones and their characteristics (ie, neutralizing or ADE). Though there was no significant difference by the χ^2 test in the percentage of cell clones producing neutralizing antibodies between the Angola- and Musoke-specific clones, the percentage of clones showing FcR-dependent ADE activity was significantly higher after immunization with Angola GP. The clones producing the enhancing antibodies were further classified into 2 groups; the first (group A) consisted of mAbs that enhanced the infectivity of both VSV-Angola and VSV-Musoke, and the second (group B) consisted of mAbs that enhanced only VSV-Angola infectivity. No mAb showed C1q-dependent ADE activity (data not shown).

Identification of ADE Epitopes on MARV GPs

To identify epitopes involved in ADE activity of mAbs, we tested the representative mAbs (IgG1) from groups A and B to determine reactivity to synthetic peptides derived from Angola and Musoke GP (PEPscreen; Sigma), and we found that all mAbs belonging to group A bound strongly to peptides corresponding to amino acid positions 330–350 and 410–430 (data not shown). We then constructed mutant GPs lacking amino acid residues 330–350 (AΔ1) or 410–430 (AΔ2) and examined the reactivity of mAbs of group A (Figure 3). In immunostaining of HEK 293T cells expressing the mutant GPs, these mAbs bound to AΔ1 but not AΔ2, suggesting that the epitopes for these mAbs lay between amino acid positions 410–430 where the sequences are partially shared between Angola and Musoke GPs. By contrast, mAbs belonging to group B failed to bind to any synthetic peptide, suggesting that their epitopes were likely in nonlinear conformation. Because mAbs of group B did not bind to Musoke GP, we constructed chimeric GPs between Angola and Musoke GPs (MA1–MA5), as indicated in Figure 3. Based on the reactivity of mAbs to the chimeric GPs expressed in HEK 293T cells, 2 regions on Angola GP were identified as the epitopes of mAbs of group B (Figure 3). mAbs AGP2-1, AGP40-32, and AGP126-10 reacted strongly with MA1 and MA2 but not MA3, MA4, and MA5, indicating that these mAbs recognized an epitope that resided between amino acid positions 369 and 385 in Angola GP. Another mAb, AGP90-28, bound to 4 chimeric GPs, MA1, MA2, MA3, and MA4, but not MA5, indicating that the epitope for this mAb was located between amino acid positions 402 and 418 in Angola GP. All of these identified epitopes were located within the MLR.

Table 1. Characterization of Monoclonal Antibody (mAb) Clones Specific to Marburg Virus Glycoprotein (GP)

	mAb Clones, no. (%)	
Clone characteristic	Angola GP	Musoke GP
ELISA positive	47 (100)	28 (100)
Neutralizing	19 (40)	9 (32)
Enhancing	15 (32) [a]	0 (0)

NOTE. Differences in the ratios of the clone numbers between Angola and Musoke GPs were analyzed by χ^2 test, and statistical significance was established at $P < .05$. ELISA, enzyme-linked immunosorbent assay.

[a] $P < .05$.

Contribution of GP2 Region to ADE

To confirm the contribution of the MLR of MARV in ADE of these mAbs, we constructed chimeric GPs whose MLRs were swapped between Angola and Musoke GPs (AMA and MAM), and the levels of infectivity of VSV pseudotyped with these chimeric GPs (VSV-AMA and VSV-MAM) were compared in K562 cells (Figure 4). Unexpectedly, however, the infectivity of VSV-MAM was not enhanced by Angola MLR-specific mAbs (ie, mAbs A2-1 and A90-28 in group B). Furthermore, in the presence of mAb AGP127-8 belonging to group A, the infectivity of VSV-AMA was still significantly higher than that of

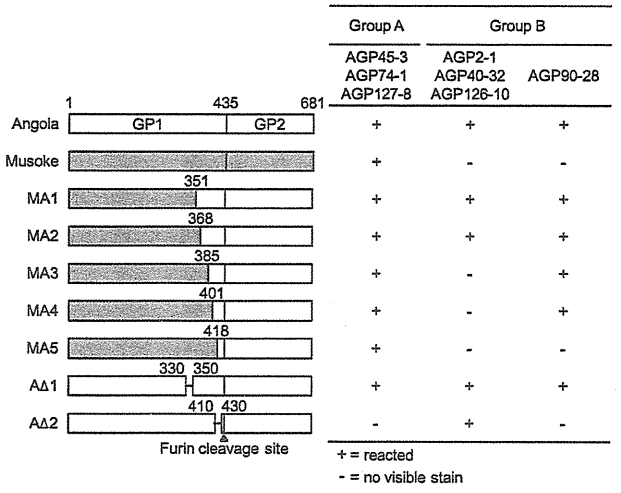

Figure 3. Reactivity of monoclonal antibodies to chimeric and deletion mutant glycoproteins (GPs). Reactivity was tested by immunostaining, as described in Materials and Methods. Numbers indicate amino acid positions of the GP primary amino acid sequence.

VSV-Musoke, and replacement of the MLR of Musoke GP with that of Angola GP (VSV-MAM) had little effect on the enhancement of infectivity compared with the infectivity of VSV-Musoke. We then examined the infectivity of VSV pseudotyped with a series of chimeric GP mutants, as shown in Figure 4, and found that only the infectivity of VSV-MAA, which had the MLR and part of GP2 (shown as GP2C in Figure 4) derived from Angola GP, was reasonably enhanced by mAbs of both groups A and B. These results indicated that both the MLR and the GP2 region of Angola GP were required for maximal ADE activity of these mAbs.

Importance of an Amino Acid at Position 547 for ADE

The GP2C regions of the Angola and Musoke GPs differ in 4 amino acids: positions 504, 547, 596, and 618. To identify which amino acid in GP2C contributed to the differential infectivity enhancement by mAbs AGP127–8, A2-1, and A90-28, the following 8 mutant GPs that contained single-amino-acid substitutions were constructed: 4 Angola-based mutant GPs (A/H504T, A/G547V, A/A596T, and A/R618K) and 4 Musoke-based mutant GPs (M/T504H, M/V547G, M/T596A, and M/K618R) (Figure 5). The infectivity of VSV pseudotyped with these mutant GPs in the presence of these mAbs was compared in K562 cells. We found that the substitution at position 547 clearly switched the viral infectivity in the presence of mAb AGP127-8, as shown by substantially enhanced and reduced infectivity of VSV-M/V547G and A/G547V, respectively, when compared with VSV pseudotyped with wild-type Musoke and Angola GPs. By contrast, the mutations at position 504, 596, and 618 had limited effects on the infectivity of the respective viruses. Interestingly, ADE of VSV-A/G547V infectivity was not observed, even in the presence of Angola GP-specific mAbs, AGP2-1 and AGP90-28. These results indicated that amino acid position 547 played an important role in the ADE activities of these mAbs. To examine whether reduced ADE activities of mAbs resulted from the loss of binding affinity due to the structural change by the mutation at position 547, we measured reactivity to each GP with ELISA (Figure 6). We found that Angola and Musoke GP antisera reacted similarly to all the GP

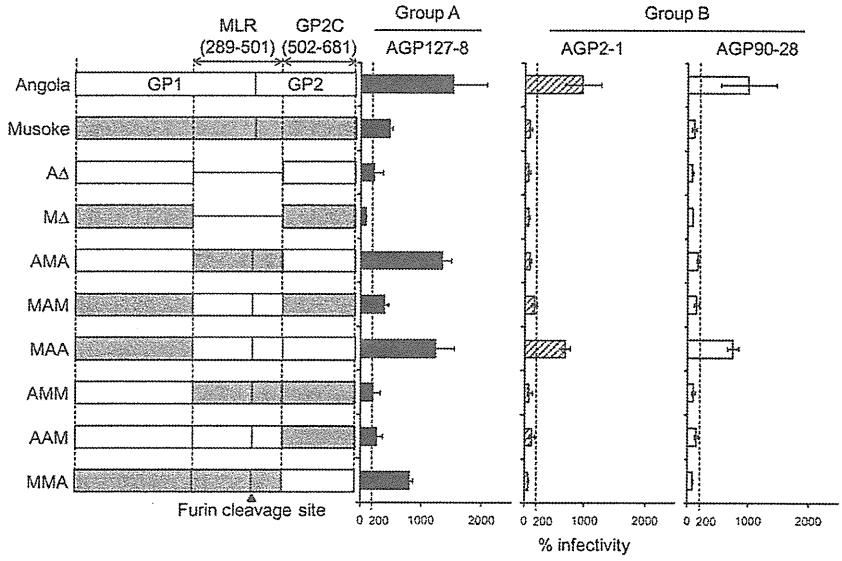

Figure 4. Infectivity of vesicular stomatitis virus pseudotyped with deletion or chimeric mutant glycoproteins (GPs) in K562 cells. VSV pseudotyped with mutant GPs was incubated with monoclonal antibodies AGP127-8, AGP2-1, and AGP90-28 (1 μg/mL) and then inoculated into K562 cells. All experiments were done in triplicate, and means and standard deviations are shown. Numbers indicate amino acid positions of the GP primary amino acid sequence.

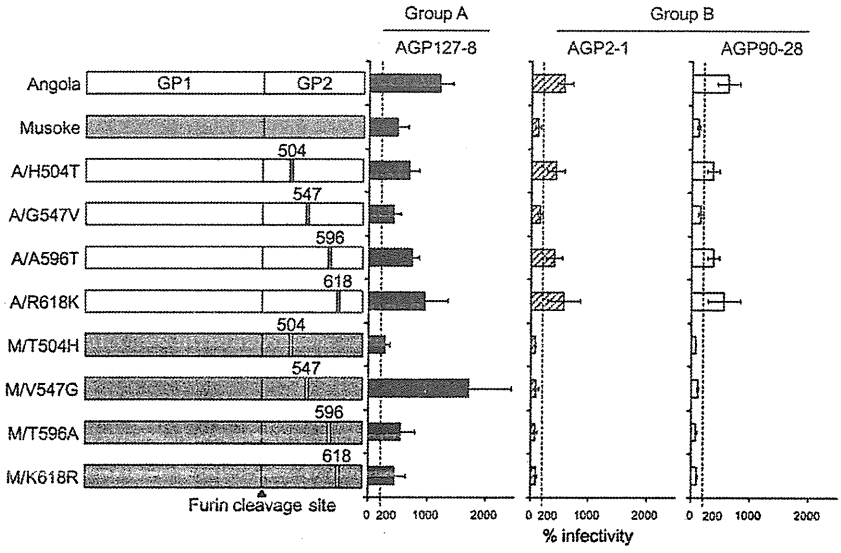

Figure 5. Infectivity of vesicular stomatitis virus pseudotyped with the mutant glycoproteins (GPs) with a single substitution in K562 cells. Experimental conditions were the same as described for Figures 1 and 4. All experiments were done in triplicate, and means and standard deviations are shown.

antigens, suggesting that the substitution at position 547 did not change the overall antigenic structure of MARV GPs. Importantly, this substitution resulted in no significant change in reactivity of these mAbs, indicating that the substitution at position 547 did not affect the binding affinity of these mAbs to GPs.

DISCUSSION

It has been shown that in addition to the common receptor- or coreceptor-dependent mechanism of cellular attachment, some viruses use antiviral antibodies for efficient entry into target cells [21]. Although this phenomenon, ADE of EBOV, has been

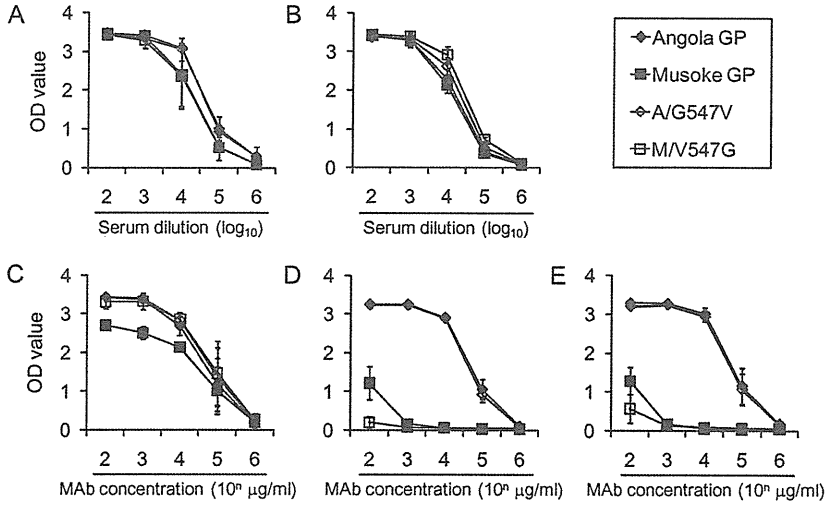

Figure 6. Binding affinities of the mouse antisera and monoclonal antibodies to Marburg virus glycoproteins (MARV GPs). Serial 10-fold dilutions of the mouse antisera to Angola GP *A* and Musoke GP *B* or indicated concentrations of mAbs AGP127-8 *C*, AGP 2-1 *D*, and AGP 90-28 *E*, were tested by enzyme-linked immunosorbent assay (ELISA) for reactivity to wild-type GPs and GPs with single amino acid substitutions. Viruslike particles expressing each GP were used as ELISA antigens. Each optical density (OD) value represents the means and standard deviations of 2 independent experiments.

studied by using mouse mAbs, convalescent human, and experimentally infected nonhuman primate serum samples [19, 20, 22], the possible contribution of MARV-specific antibodies to ADE has not been reported previously to our knowledge.

In the present study, we demonstrated that the infectivity of VSV-Angola in K562 cells was enhanced notably in the presence of Angola GP antisera (ie, FcR-dependent ADE), whereas Musoke GP antisera did not significantly enhance the infectivity of VSV-Angola or -Musoke (Figure 1). This difference between the 2 MARV strains was also supported by the observation that immunization with Angola GP induced significantly higher numbers of B-cell clones with infectivity-enhancing properties than did immunization with Musoke GP (Table 1). These results may suggest that the potential difference in the pathogenicity between the MARV strains Angola and Musoke might be partially explained by the ability to induce infectivity-enhancing antibodies, as was proposed for the distinct pathogenicity seen with ZEBOV and Reston EBOV [20, 22].

However, it has been argued whether the strain Angola is more virulent for humans and nonhuman primates than the strain Musoke. Along with the high case fatality rate in the Angola epidemic, it was also noted that in most earlier outbreaks of Marburg hemorrhagic fever, the case fatality rates did not exceed 50% [2, 4, 5], but there was another outbreak in which the case fatality rate was 83%, quite similar to what was seen in the Angola outbreak [29]. In experimental infection of macaques, 3 animals infected with the strain Angola died after illness that was more rapidly progressive than those caused by other viruses tested in other experiments [8, 9, 30], and necropsy showed infection of the liver with accompanying necrosis. However, because animals infected with the strains Angola and Musoke have not been compared in the same experiment, it could only be concluded that infection with the strain Angola "appeared to progress more rapidly" than other MARVs tested in other studies. In addition, one of the patients infected with the strain Musoke developed a rapidly progressive illness that closely resembled the features of illness seen in macaques experimentally infected with the strain Angola [5, 9]. Thus, although our results may suggest the possible contribution of ADE to differences in pathogenicity between the strains Angola and Musoke, further investigations are needed to conclude that the strain Angola has a uniquely higher pathogenicity.

Compared with ZEBOV GP antisera, Angola GP antisera enhanced the infectivity of VSV-Angola less efficiently in the presence of C1q (Figure 2*A*). We therefore compared the profile of IgG subclasses in the antisera. A significant difference was seen in the levels of IgG2a, IgG2b, and IgG3 between MARV Angola and ZEBOV GP antisera (Figure 2*B*). It is likely that IgG2a and IgG3 play a more prominent role in C1q-dependent ADE activity, because these classes of antibodies are thought to have a higher affinity for C1q molecules than IgG1 [20]. Thus, this difference may account for the lower activity of C1q-dependent ADE in Angola antisera. It will be of interest to elucidate why MARV and EBOV GPs induce a distinct antibody repertoire.

It was shown that most of the ADE epitopes of EBOV GP were located in the MLR [20]. In the present study, we identified 3 distinct ADE epitopes on MARV GP, and these ADE epitopes were mostly located in the MLR of the GP1 subunit (Figure 3). To further ascertain the contribution of the MLR of MARV GP to ADE, we examined the infectivity of VSV pseudotyped with the Angola GP lacking the MLR in K562 cells in the presence of the Angola GP antisera and found that ADE activity was reduced drastically (data not shown). These results suggested that the MLR of MARV GP also contained dominant epitopes for antibodies that enhance viral infectivity.

Two ADE epitopes, amino acid positions 369–385 and 402–418 (group B epitopes), were recognized by Angola GP-specific mAbs, whereas another epitope, amino acids 410–430, was shared between the 2 GPs (group A epitope). Accordingly, mAbs belonging to group A bound to both Angola and Musoke GP and enhanced the infectivity of both VSV-Angola and -Musoke (Figures 4 and 5). These results suggest that Angola and Musoke GPs share some epitopes recognized by the ADE antibodies. However, Musoke GP antisera and the Musoke GP-specific mAbs did not show ADE activity, suggesting that Musoke GP contained fewer ADE epitopes than Angola GP. Thus, we propose that the primary amino acid structure and glycosylation pattern of the MLR may influence epitope exposure on the GP molecule.

It was noted that the substitution at amino acid position 547 in the GP2 region affected ADE significantly but did not change the binding affinity of the mAbs to the Angola and Musoke GPs (Figure 6). Amino acid position 547 is located at the base of the internal fusion loop [24]. We demonstrated elsewhere that a single-amino acid substitution at position 547 affected the efficiency of C-type lectin-mediated entry and the mechanism underlying endosomal entry such as proteolytic processing by endosomal cathepsin [24]. The present findings also suggest that the amino acid at position 547 plays a role in entry mediated through the ADE pathway. When virus particles are internalized in endosomes via the ADE pathway, it is possible that (1) glycine at position 547 weakens the interaction between GP1 and GP2, resulting in increased cathepsin susceptibility or reduced cathepsin dependence in endosomes and (2) the amino acid at position 547 affects the flexibility of the fusion loop and/or the conformational change needed for fusion. It would be interesting to investigate the role of this particular amino acid for proper GP functions.

In this study, we demonstrated in vitro ADE for MARV that might be associated with the distinct pathogenicity of certain MARV strains. The ADE epitopes were localized in the MLR of GP. Furthermore, our data suggest that the efficiency of ADE-mediated entry of MARV is controlled by the epitope structure on GP and the mechanisms underlying endosomal

entry. Further in vitro studies are required to prove the contribution of ADE to the exacerbation of MARV infection; these studies should use serum samples from infected monkeys and patients to confirm whether ADE is induced by actual MARV infection and include animal experiments using a reverse-genetics system to introduce mutations into the ADE epitopes and reduce the production of ADE antibodies. The present study, however, together with our previous study [24], provides new insights into the molecular basis of MARV entry, mediated by interaction between the GP MLR and cellular attachment factors, such as C-type lectins and MLR-specific antibodies.

Funding

This work was supported by Research Fellowships for Young Scientists from the Japan Society for the Promotion of Science, the Takeda Science Foundation, a Grant-in-Aid for Scientific Research on Priority Areas (grant 19041001), and, in part, by the Program of Founding Research Centers for Emerging and Reemerging Infectious Diseases (grant 05021011) and the Global COE Program "Establishment of International Collaboration Centers for Zoonosis Control" (grant F-001) from the Ministry of Education, Culture, Sports, Science and Technology, Japan (http://www.mext.go.jp/english/index.htm). Filovirus work at the Rocky Mountain Laboratories is funded by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Acknowledgments

We thank Hiroko Miyamoto, Ayaka Yokoyama, Teiji Murakami, and Aiko Ohnuma for technical assistance and Kim Barrymore for editing the manuscript.

References

1. Sanchez A, Geisbert TW, Feldmann H. Filoviridae: Marburg and Ebola viruses. In Knipe DM, Howley PM, Griffin DE eds. Fields virology, 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins, **2006**; 1409–48.
2. Martini G. Marburg virus disease. Postgrad Med J **1973**; 49:542–6.
3. Johnson E, Johnson B, Silverstein D, et al. Characterization of a new Marburg virus isolated from a 1987 fatal case in Kenya. Arch Virol Suppl **1996**; 11:101–14.
4. Gear J, Cassel G, Gear A, et al. Outbreak of Marburg virus disease in Johannesburg. Br Med J **1975**; 4:489–93.
5. Smith D, Johnson B, Isaacson M, et al. Marburg-virus disease in Kenya. Lancet **1982**; 1:816–20.
6. Bausch D, Borchert M, Grein T, et al. Risk factors for Marburg hemorrhagic fever, Democratic Republic of the Congo. Emerg Infect Dis **2003**; 9:1531–7.
7. Towner J, Khristova M, Sealy T, et al. Marburgvirus genomics and association with a large hemorrhagic fever outbreak in Angola. J Virol **2006**; 80:6497–516.
8. Daddario-DiCaprio K, Geisbert T, Geisbert J, et al. Cross-protection against Marburg virus strains by using a live, attenuated recombinant vaccine. J Virol **2006**; 80:9659–6.
9. Geisbert T, Daddario-DiCaprio K, Geisbert J, et al. Marburg virus Angola infection of rhesus macaques: pathogenesis and treatment with recombinant nematode anticoagulant protein c2. J Infect Dis **2007**; 196(Suppl 2):S372–81.
10. Centers for Disease Control and Prevention. Update: filovirus infection in animal handlers. MMWR Morb Mortal Wkly Rep **1990**; 39:221.
11. Fisher-Hoch S, McCormick J. Experimental filovirus infections. Curr Top Microbiol Immunol **1999**; 235:117–43.
12. Dube D, Brecher M, Delos S, et al. The primed ebolavirus glycoprotein (19-kilodalton GP1,2): sequence and residues critical for host cell binding. J Virol **2009**; 83:2883–91.
13. Kuhn J, Radoshitzky S, Guth A, et al. Conserved receptor-binding domains of Lake Victoria marburgvirus and Zaire ebolavirus bind a common receptor. J Biol Chem **2006**; 281:15951–8.
14. Ito H, Watanabe S, Sanchez A, Whitt M, Kawaoka Y. Mutational analysis of the putative fusion domain of Ebola virus glycoprotein. J Virol **1999**; 73:8907–12.
15. Manicassamy B, Wang J, Rumschlag E, et al. Characterization of Marburg virus glycoprotein in viral entry. Virology **2007**; 358:79–88.
16. Yang Z, Duckers H, Sullivan N, Sanchez A, Nabel E, Nabel G. Identification of the Ebola virus glycoprotein as the main viral determinant of vascular cell cytotoxicity and injury. Nat Med **2000**; 6:886–9.
17. Sanchez A, Trappier S, Mahy B, Peters C, Nichol S. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. Proc Natl Acad Sci U S A **1996**; 93:3602–7.
18. Sanchez A, Trappier S, Stroher U, Nichol S, Bowen M, Feldmann H. Variation in the glycoprotein and VP35 genes of Marburg virus strains. Virology **1998**; 240:138–46.
19. Takada A, Feldmann H, Ksiazek TG, Kawaoka Y. Antibody-dependent enhancement of Ebola virus infection. J Virol **2003**; 77:7539–44.
20. Takada A, Ebihara H, Feldmann H, Geisbert T, Kawaoka Y. Epitopes required for antibody-dependent enhancement of Ebola virus infection. J Infect Dis **2007**; 196(Suppl 2):S347–56.
21. Takada A, Kawaoka Y. Antibody-dependent enhancement of viral infection: molecular mechanisms and in vivo implications. Rev Med Virol **2003**; 13:387–98.
22. Takada A, Watanabe S, Okazaki K, Kida H, Kawaoka Y. Infectivity-enhancing antibodies to Ebola virus glycoprotein. J Virol **2001**; 75:2324–30.
23. Takada A, Robison C, Goto H, et al. A system for functional analysis of Ebola virus glycoprotein. Proc Natl Acad Sci U S A **1997**; 94: 14764–9.
24. Matsuno K, Kishida N, Usami K, et al. Different potential of C-type lectin-mediated entry between Marburg virus strains. J Virol **2010**; 84:5140–7.
25. Noda T, Sagara H, Suzuki E, Takada A, Kida H, Kawaoka Y. Ebola virus VP40 drives the formation of virus-like filamentous particles along with GP. J Infect Dis **2002**; 76:4855–65.
26. Swenson D, Warfield K, Kuehl K, et al. Generation of Marburg virus-like particles by co-expression of glycoprotein and matrix protein. FEMS Immunol Med Microbiol **2004**; 40:27–31.
27. Shahhosseini S, Das D, Qiu X, Feldmann H, Jones SM, Suresh MR. Production and characterization of monoclonal antibodies against different epitopes of Ebola virus antigens. J Virol Methods **2007**; 143:29–37.
28. Nakayama E, Yokoyama A, Miyamoto H, et al. Enzyme-linked immunosorbent assay for the detection of filovirus species-specific antibodies. Clin Vaccine Immunol **2010**; 17:1723–8.
29. Bausch DG, Nichol ST, Muyembe-Tamfum JJ, et al. Marburg hemorrhagic fever associated with multiple genetic lineages of virus. N Engl J Med **2006**; 355:909–19.
30. Daddario-DiCaprio KM, Geisbert TW, Ströher U, et al. Postexposure protection against Marburg haemorrhagic fever with recombinant vesicular stomatitis virus vectors in non-human primates: an efficacy assessment. Lancet **2006**; 367:1399–404.

OPEN ACCESS Freely available online

PLoS one

Gnarled-Trunk Evolutionary Model of Influenza A Virus Hemagglutinin

Kimihito Ito[1,2], Manabu Igarashi[1], Yutaka Miyazaki[3], Teiji Murakami[1], Syaka Iida[1], Hiroshi Kida[1,4,5,6], Ayato Takada[1,7]*

1 Hokkaido University Research Center for Zoonosis Control, Sapporo, Japan, 2 PRESTO, Japan Science and Technology Agency, Saitama, Japan, 3 Faculty of Liberal Arts and Sciences, Osaka University of Economics and Law, Yao, Japan, 4 Department of Disease Control, Graduate School of Veterinary Medicine, Hokkaido University, Sapporo, Japan, 5 OIE Reference Laboratory for Highly Pathogenic Avian Influenza, Sapporo, Japan, 6 SORST, Japan Science and Technology Agency, Saitama, Japan, 7 School of Veterinary Medicine, The University of Zambia, Lusaka, Zambia

Abstract

Human influenza A viruses undergo antigenic changes with gradual accumulation of amino acid substitutions on the hemagglutinin (HA) molecule. A strong antigenic mismatch between vaccine and epidemic strains often requires the replacement of influenza vaccines worldwide. To establish a practical model enabling us to predict the future direction of the influenza virus evolution, relative distances of amino acid sequences among past epidemic strains were analyzed by multidimensional scaling (MDS). We found that human influenza viruses have evolved along a gnarled evolutionary pathway with an approximately constant curvature in the MDS-constructed 3D space. The gnarled pathway indicated that evolution on the trunk favored multiple substitutions at the same amino acid positions on HA. The constant curvature was reasonably explained by assuming that the rate of amino acid substitutions varied from one position to another according to a gamma distribution. Furthermore, we utilized the estimated parameters of the gamma distribution to predict the amino acid substitutions on HA in subsequent years. Retrospective prediction tests for 12 years from 1997 to 2009 showed that 70% of actual amino acid substitutions were correctly predicted, and that 45% of predicted amino acid substitutions have been actually observed. Although it remains unsolved how to predict the exact timing of antigenic changes, the present results suggest that our model may have the potential to recognize emerging epidemic strains.

Citation: Ito K, Igarashi M, Miyazaki Y, Murakami T, Iida S, et al. (2011) Gnarled-Trunk Evolutionary Model of Influenza A Virus Hemagglutinin. PLoS ONE 6(10): e25953. doi:10.1371/journal.pone.0025953

Editor: Art F. Y. Poon, British Columbia Centre for Excellence in HIV/AIDS, Canada

Received June 7, 2011; **Accepted** September 13, 2011; **Published** October 10, 2011

Funding: This project was supported by the Program of Founding Research Centers for Emerging and Reemerging Infectious Diseases, the Japan Initiative for Global Research Network on Infectious Diseases (J-GRID), the Global COE Program, Grants-in-Aid, all from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, and PRESTO and SORST from Japan Science and Technology Agency (JST) Basic Research Programs. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: atakada@czc.hokudai.ac.jp

Introduction

The hemagglutinin (HA) molecule of influenza A viruses is the prime target of antibodies that neutralize viral infectivity. The strong immune pressure against HA in the human population selects a new variant every 2–5 years [1–6]. Thus influenza A viruses undergo antigenic changes with gradual accumulation of amino acid substitutions on HA, and the antigenic change is one of the primary reasons why vaccination is not a perfect measure to control seasonal influenza. Accordingly, influenza vaccine often requires replacement to avoid antigenic mismatch between vaccine and epidemic strains [7]. The decision of vaccine replacement must be made several months before a minor strain become dominant strain [8]. Thus the prediction of antigenic change of influenza A virus [2,9–15] has been one of the major public health goals.

Phylogenetic analyses of HA genes of human H3N2 viruses have revealed the presence of a long main trunk and short side branches in their evolutionary tree. The main trunk has grown continuously from a pandemic strain in 1968 to recent epidemic strains, and tips of each branch reached a dead end on the evolutionary pathway [13,14,16–18]. This 'cactus-like' phylogenetic tree indicates that the viruses on the side branches do not produce next epidemic strains, while the viruses near the main trunk do contribute to the production of both an epidemic strain and its next epidemic strain. Although two or more antigenically different strains were known to co-circulate in a single epidemic season [13,19,20], the single-trunk phylogenetic tree indicates the diversity of the HA amino acid sequences at any point in time is relatively limited. The reason why only one trunk exists has yet to be fully understood, but several theories have been proposed to explain this phenomenon [21–23].

The aims of the present studies are to establish a practical model enabling us to predict the evolutionary direction of the virus that causes future epidemics and to examine the accuracy of the prediction based on the model. First we analysed relative distances of amino acid sequences among past epidemic strains using a method called multidimensional scaling (MDS) [24]. We found that human influenza viruses have evolved along a gnarled evolutionary pathway with an approximately constant curvature in the MDS-constructed 3D space. The constant curvature was reasonably explained by assuming that the rate of amino acid substitutions varied from one position to another according to a gamma distribution. The estimated parameters of the gamma

distribution allowed us to predict the amino acid substitutions on HA in subsequent years with reasonable accuracy, indicating the potential to select suitable vaccine strains for the subsequent epidemic seasons.

Results

To expose underlying patterns of HA amino acid substitutions in the evolutionary pathway along the main trunk, we conducted multidimensional scaling (MDS) analysis [24] of HA sequences. The fundamental idea for visualizing a large number of sequences in a low dimensional space is based on the same idea described in a recent paper by He and Deem [15]. By performing MDS analysis, one may obtain a visual map of objects where the dissimilarity between objects is represented as the distance between corresponding points. A total of 2,640 unique amino acid sequences of the HA1 [3,25] domain (328 amino acids long) of the H3N2 viruses isolated from humans during the period from 1968 through 2009 were analysed by MDS and visualized in a three-dimensional (3D) space (Figure 1, Movie S1). In the resulting 3D map, each HA sequence was represented as a point, and the number of different amino acids between two HA sequences was represented as the relative distance between two corresponding points. Although the original amino acid sequences provided 328 dimensional data, the numbers of different amino acids among sequences were reasonably approximated by distance in this 3-dimensional map with a root-mean-square error of 1.72 (Figure S1). Consistent with phylogenetic analyses, viruses that were isolated close in time were located near each other, forming a thick main trunk with short branches elongated from the trunk (Figure 1A, Figure 1B). The main trunk grew continuously from a pandemic strain in 1968 to recent epidemic strains. Each branch consisted of epidemic strains isolated during a period of 3–5 years.

It should be noted that the MDS representation revealed a characteristic feature that has not been clear only from phylogenetic trees, the observation of a gnarled trunk constantly curved in the 3D map. Since amino acid substitutions on human virus HA mainly occur in the HA1 domain [3], this result indicated that the evolution of H3N2 virus HA was characterized by this gnarled evolutionary pathway.

We also conducted the same analysis for H1N1 human viruses (Figure S2, Movie S2). The 3D representation of HA sequences of H1N1 seasonal influenza viruses showed the same pattern in their evolutionary pathways, a long gnarled trunk elongated from the pandemic strain in 1918. Another H1N1 pandemic strain, which was introduced into the human population in 2009 [26], was located at the end of a step-wise path that consisted of swine H1N1 influenza viruses isolated from humans from 1976 to 2007 [27].

In the 3D map in Figure 1, the spatial distance between each pair of sequences represents the number of different amino acids between these sequences. Figure 2 illustrates two distinct patterns of amino acid substitutions that produce different spatial arrangements of viruses in this map. If a series of amino acid substitutions all occur in different positions, then the distance from an ancestor to a mutant should be proportional to the number of substitutions. These independent substitutions make a straight arrangement of viruses on the map (Figure 2A, 2B). On the other hand, if amino acids at particular positions were substituted more than once, the distance from an ancestor to a mutant should be less than the number of substitutions. These concentrated substitutions at the same amino acid positions make a curved arrangement of viruses (Figure 2C, 2D). For this reason, the gnarled trunk found in the MDS representation of HA sequences (Figure 1) indicated that HA variants on the trunk favoured multiple amino acid substitutions at the same positions. The fact that most of the amino acid substitutions occurred near antigenic domains A–E [3,5,9,28] was consistent with the observation of the curved trunk in the 3D map.

To investigate the property of the curvature of the main trunk, we analysed the distribution of the number of positions that were substituted on the trunk from 1968 to 2009 (Figure 3A, Table S1). Of the 328 positions on the HA1 sequence, 260 remained unchanged for 41 years. At 36 amino acid positions, residues were substituted once, and at 19 positions twice. The number of positions gradually decreased as their observed frequency of substitutions increased, but there was one position that has substituted eight times. The mean of the substitution frequencies was 0.384, and the variance of the substitution frequencies was 0.904. Given these statistics for the substitution frequency, the shape of the histogram is almost identical to the curve of a gamma distribution [29] having the same mean and variance (Figure 3A). From this result, it is likely that the rate of amino acid substitutions varies from one position to another according to a gamma distribution.

To estimate the parameters of the gamma distribution precisely, non-linear regression analysis was performed. First, 91 HA sequences near the main trunk were selected as trunk sequences. In Figure 3B, the number of different amino acids between two trunk sequences (Figure 3C) is plotted against the difference in their isolation years. It is known that if the amino acid substitution rate varies according to a gamma distribution, the expected number of different amino acids between two sequences can be calculated by the formula: $d = L(1-(a/(a+\bar{r}t))^a)$, where $L$is the length of the sequences, a is the shape parameter of the gamma distribution (gamma parameter), $\bar{r}$ is the mean substitution rate, and t is the difference in the years of the two sequences [30]. By fitting the above formula to the actual numbers of different amino acids on the trunk, the gamma parameter and mean substitution rate were estimated ($a = 0.129$ and $\bar{r} = 0.0118$), showing a good fit to the actual data ($P<0.001$). This result indicated that the number of different amino acids between two sequences could be determined from the difference of the year of isolation, according to the gamma-distribution-based model presented above. Thus, it is reasonable to conclude that the constant curvature on the trunk in the MDS representation (Figure 1) was attributed to this nature.

Next, an attempt was made to apply the gamma-distribution-based substitution model of the trunk to the prediction of amino acid substitutions in subsequent years. The key idea of our prediction method is to select the direct progenitor virus for future epidemics from the surveillance samples of each year. We designate a virus strain that is located near the trunk extending into the next year as a Leading Bud. A Leading Bud can be considered as an potential dominant strain that is not dominant in the given year and become dominant the next year [15]. Using the formula $d = 328 \times (1-(a/(a+\bar{r}t))^a)$under the estimated $a$and $\bar{r}$, one may calculate the expected number of different amino acids between two HA sequences located on the trunk. According to this gamma-distribution-based formula, a virus that will appear on the trunk in a particular year is expected to have 4, 7, 10, 13, ... , 60 different amino acids in HA, when compared with viruses in 1, 2, 3, 4, ... , 42 years before, respectively. Therefore, given the large variety of viruses isolated in a year, the virus that is likely to be located near the extended trunk is the virus to which HA sequence dissimilarities from past viruses have the highest fit to those expected under the gamma-distribution-based formula (Figure 3D).

To examine whether the method correctly selected the Leading Bud in the subsequent year, we conducted retrospective tests for each year from 1997 to 2009. Evaluation was made by comparing

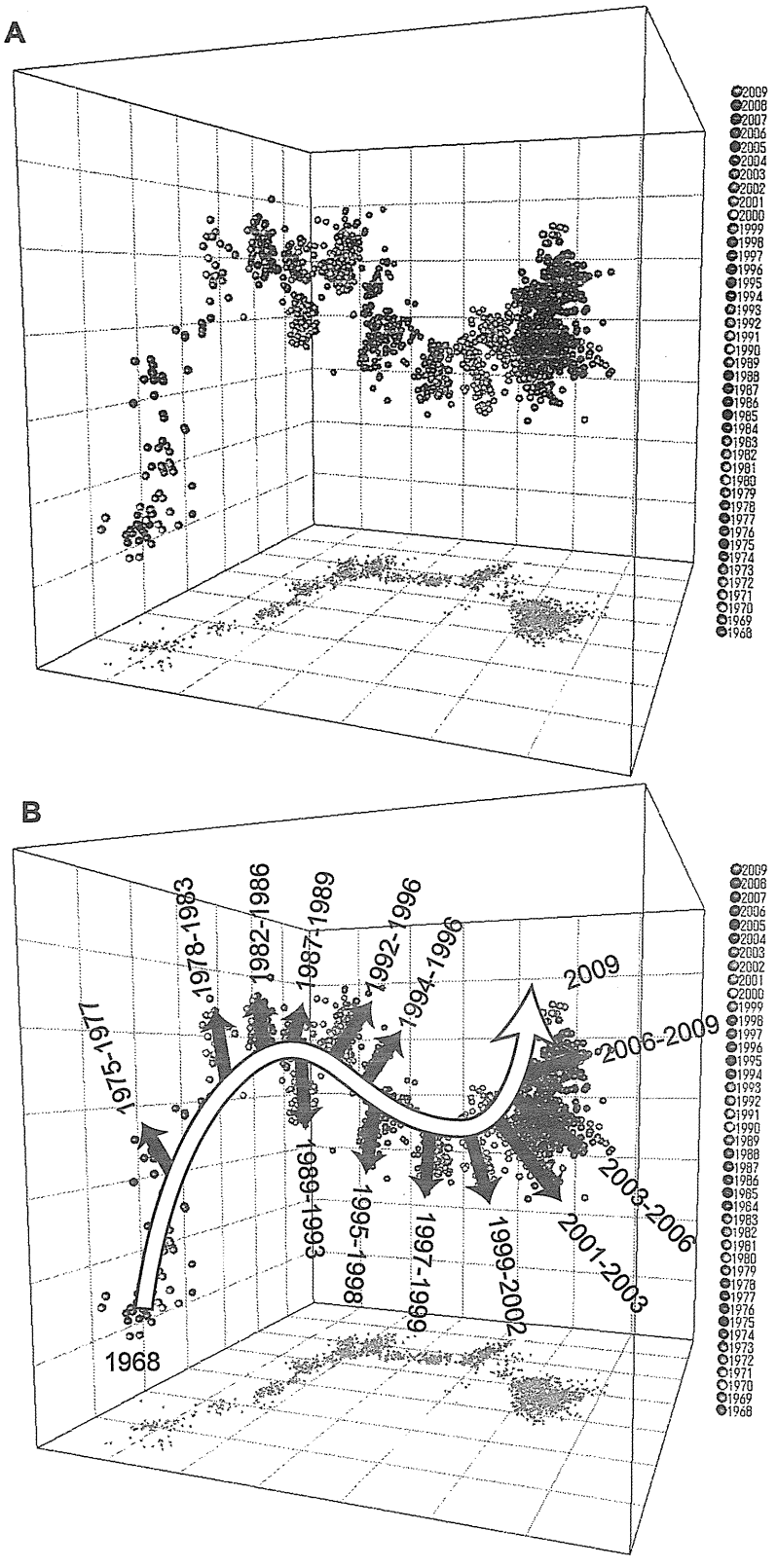
A
B
2009
2008
2007
2006
2005
2004
2003
2002
2001
2000
1999
1998
1997
1996
1995
1994
1993
1992
1991
1990
1989
1988
1987
1986
1985
1984
1983
1982
1981
1980
1979
1978
1977
1976
1975
1974
1973
1972
1971
1970
1969
1968
1968
1975-1977
1978-1983
1982-1986
1987-1989
1989-1993
1992-1996
1994-1996
1995-1998
1997-1999
1999-2002
2001-2003
2003-2006
2006-2009
2009

Figure 1. Three-dimensional map of HA sequences of H3N2 human influenza A viruses. A total of 2,640 amino acid sequences of the HA1 domain of human H3N2 influenza A viruses isolated during the period from 1968 through 2009 are visualized in a 3D space. Each point represents a virus strain. The distance between two viruses in the 3D map represents the number of different amino acids between their HA sequences. The whole coordination is determined by MDS analysis. The root-mean-square error of the 3D map was 1.72. All three axes represent the sequence dissimilarity (spacing between grid lines represents 10 different amino acids), and the configuration can be freely rotated and translated. Shadows represent projections of points onto the coordinate planes. (A) the 3D map colour-coded by the year of isolation of the virus. (B) a schematic diagram of the 3D map.
doi:10.1371/journal.pone.0025953.g001

the predicted Leading Bud with the dominant sequence of the subsequent year, which consisted of the amino acids that constituted the majority at each position in the year. Table 1 shows the results of retrospective tests (for details, see Table S2). The recall, which is the probability that an actual substitution was correctly predicted, was 1.00 in 4 of the 12 calendar years. The overall recall of the prediction was 0.70, indicating that the model had a reasonable ability to predict amino acid substitution in the subsequent year for each year. The precision, which is the probability that a predicted substitution actually occurred, varied from 0.0 to 0.89, and the overall recall of the prediction was 0.45.

To assess the validity of the result of the retrospective tests, we repeated similar retrospective tests with other methods and compared the results (Table 2). First we tested a method that randomly selects an HA sequence for each year. With 100 sets of tests, overall precision and recall were 0.24 ± 0.015 and 0.22 ± 0.012 respectively, showing low predictive ability as expected. Secondly we tested method that selects the HA sequence that has the maximum numbers of substitutions at the 18 positively selected codons identified by Bush et al [10]. Although the overall precision and recall were much higher than random tests, the accuracy of prediction was lower than that of our method. Three methods that select the sequence that has the maximum number of amino acid substitutions from current and past dominant sequences yielded higher recalls. However, the overall precision was lower than the method using 18 positively selected codons. Among all the methods we tested, the gamma-distribution-based method was the only method that yields higher recall and higher precision than Bush's methods.

Discussion

Our study found that the long-term evolution of HA was reasonably characterised by a 'bonsai-like' pathway of which trunk

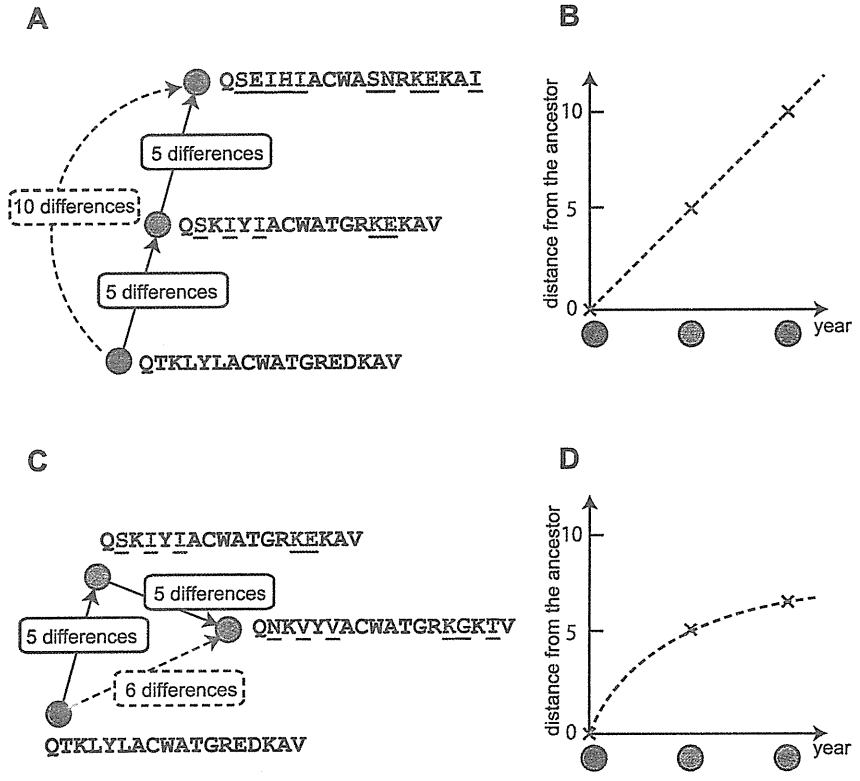

Figure 2. Two distinct patterns of amino acid substitutions that produce different spatial arrangements of viruses. (A) The straight arrangement of viruses. If a series of amino acid substitutions all occur in different positions, then the distance from an ancestor to a mutant should be proportional to the number of substitutions (B). These independent substitutions make a straight arrangement of viruses on the map. (C) The curved arrangement of viruses. If amino acids at particular positions were substituted more than once, the distance from an ancestor to a mutant should be less than the number of substitutions (D). These concentrated substitutions at the same amino acid positions make a curved arrangement of viruses. In both panels, viruses are represented by circles, with illustrative examples of their amino acid sequences and substitutions on them.
doi:10.1371/journal.pone.0025953.g002