# Homologous Recombination as an Evolutionary Force in the Avian Influenza A Virus

*Cheng-Qiang He,* Zhi-Xun Xie,† Guan-Zhu Han,* Jian-Bao Dong,† Dong Wang,* Jia-Bo Liu,† Le-Yuan Ma,* Xiao-Fei Tang,† Xi-Ping Liu,* Yao-Shan Pang,† and Guo-Rong Li**

*Department of Biotechnology College of Life Science, Shandong Normal University, Shandong Province, Jinan, China; and †Department of Biotechnology Guangxi Veterinary Research Institute, Nanning, China

Avian influenza A viruses (AIVs), including the H5N1, H9N2, and H7N7 subtypes, have been directly transmitted to humans, raising concerns over the possibility of a new influenza pandemic. To prevent a future avian influenza pandemic, it is very important to fully understand the molecular basis driving the change in AIV virulence and host tropism. Although virulent variants of other viruses have been generated by homologous recombination, the occurrence of homologous recombination within AIV segments is controversial and far from proven. This study reports three circulating H9N2 AIVs with similar mosaic *PA* genes descended from H9N2 and H5N1. Additionally, many homologous recombinants are also found deposited in GenBank. Recombination events can occur in PB2, PB1, PA, HA, and NP segments and between lineages of the same/different serotype. These results collectively demonstrate that intragenic recombination plays a role in driving the evolution of AIVs, potentially resulting in effects on AIV virulence and host tropism changes.

## Introduction

Recently, a highly pathogenic avian influenza A virus (H5N1) has resulted in the deaths of more than 200 people (http://www.who.int/csr/disease/avian_influenza/country/cases_table_2008_09_10/en/index.html) and millions of poultry in Asia, Europe, and Africa (Chang et al. 2007). In addition to H5N1, purely avian influenza viruses (AIVs), specifically the H9N2 and H7N7 subtypes, have been directly transmitted to humans, raising concerns over the possibility of a new influenza pandemic among the world's immunologically naive populations (Horimoto and Kawaoka 2005). In order to prevent or limit a future avian influenza pandemic, it is very important to fully understand the molecular basis driving the change in AIV virulence and host tropism.

Until now, re-assortment and dynamic gene mutation have been considered the key factors responsible for the evolution of AIV and have thus been studied in detail (Nelson and Holmes 2007). It has been previously reported that a re-assortment of gene segments was responsible for creating an entirely novel influenza A virus strain capable of infecting humans (Steinhauer and Skehel 2002). However, phylogenetic analyses of amino acid changes suggested that avian influenza viruses, unlike mammalian strains, show low evolutionary rates (Gorman et al. 1990); thus, dynamic gene mutation must still play an important role in the virulence change of AIV (Hatta et al. 2001). In addition, it has been shown that homologous recombination can also play an important role in the evolution of some positive-strand RNA viruses (Kirkegaard and Baltimore 1986; Nagy and Simon 1997; Peiris et al. 1999). Virulent variants of some other viruses have been generated by homologous recombination (Worobey et al. 1999; Anderson et al. 2000; Pita et al. 2001). Even recombination between a vaccine strain and a persisting pestivirus resulted in a cytopathogenic virus and induction of lethal disease (Becher et al. 2001). Although there is evidence that influenza viruses undergo various forms of nonhomologous recombination (Khatchikian et al. 1989; Orlich et al. 1994), the occurrence of homologous recombination within segments is highly controversial. Gibbs et al. (2001) have proposed that recombination occurred in the *HA* gene between human and swine influenza viruses. However, the evidence of recombination has been questioned because of the absence of phylogenetic support (Worobey et al. 2002). Interestingly, it has recently been reported that human influenza viruses do not recombine (Boni et al. 2008). Therefore, there is much evidence still needed to demonstrate that homologous recombination occurs within segments.

In this study, we report on three circulating H9N2 AIVs with mosaic gene(s) isolated in China in the year 2000. To determine whether homologous recombination really shapes the evolution of AIVs and to provide some insights into the recombination itself, we analyzed roughly 9,000 complete segments of AIVs deposited in GenBank and found the evidence of recombination in these genes. Moreover, several mosaic viruses were also found to be able to circulate in the field. These data demonstrate that intragenic recombination can act to drive the evolution of AIVs.

## Materials and Methods

### Viruses

Viruses of the H9N2 subtype, A/chicken/Guangxi/1/00, A/chicken/Guangxi/14/00, and A/chicken/Guangxi/17/00, were isolated from dead chickens exhibiting various clinical symptoms of avian influenza illness in three counties of the Guangxi Province of China in 2000. A/chicken/Guangxi/14/00 and A/chicken/Guangxi/17/00 were collected in November, whereas A/chicken/Guangxi/1/00 was collected in March. The serotype of each isolate was determined by hemagglutination inhibition and neuraminidase inhibition tests using polyclonal chicken antisera, standard hemagglutination inhibition antisera (bought from the Harbin Veterinary Research Institute of China: Harbin City, China) and neuraminidase inhibition antisera (a generous gift from Dr Shortridge of Hong Kong University).

Key words: Avain influenza virus, homologous recombination, evolution.

E-mail: hchqiang@yahoo.com.cn.

*Mol. Biol. Evol.* 26(1):177–187. 2009
doi:10.1093/molbev/msn238
Advance Access publication October 17, 2008

**Table 1**
**The Number of Complete Gene Sequences Analyzed and the Number of Mosaic Segments Found in This Study**

| | H5N1 | H9N2 | H6N1 | H7N1 | H7N2 | H7N7 | H6N2 | H7N3 | H4N6 | H3N8 | Total | Mosaics |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PB2 | 382 | 94 | 52 | 21 | 5 | 13 | 50 | 38 | 46 | 43 | 744 | 13 |
| PB1 | 397 | 95 | 53 | 20 | 5 | 11 | 50 | 39 | 46 | 48 | 764 | 8 |
| PA | 472 | 113 | 52 | 23 | 4 | 12 | 45 | 40 | 42 | 42 | 845 | 14 |
| HA | 841 | 156 | 60 | 51 | 26 | 54 | 71 | 62 | 53 | 97 | 1,471 | 4 |
| NP | 496 | 136 | 56 | 29 | 4 | 19 | 54 | 43 | 61 | 53 | 951 | 5 |
| NA | 839 | 286 | 65 | 31 | 54 | 15 | 69 | 64 | 58 | 78 | 1,559 | 1 |
| MP | 558 | 204 | 55 | 26 | 52 | 21 | 64 | 60 | 125 | 74 | 1,239 | 0 |
| NS | 615 | 292 | 60 | 24 | 57 | 26 | 79 | 58 | 141 | 74 | 1,426 | 0 |
| Mosaic | Yes | Yes | No | No | No | No | No | No | No | No | No | |
| Total | 4,600 | 1,376 | 453 | 225 | 207 | 171 | 482 | 404 | 572 | 509 | 8,999 | 45 |

The three isolates were determined to be of the H9N2 subtype. Before further studies were conducted, the viruses were purified using a plaque-forming method in chicken embryo fibroblast cells. Purified viruses were cultured in 10-day-old, specific pathogen–free embryonated chicken eggs for sequence analysis.

### Sequencing

Viral RNA was extracted from allantoic fluid using a protocol described previously (Xie et al. 1997). Amplification of the eight full-length genes was carried out by reverse transcription polymerase chain reaction using multipairs of specific primers listed in table 4. DNA fragments were cloned into the pMD18-T easy vector (Takara Biotechnology, Dalian, China) and sequenced by Takara Biotechnology. These sequences have been deposited in GenBank (Accession No. DQ485205–DQ485228).

### Analysis of Recombination

The reference sequences for recombination analysis were downloaded from GenBank. Multialignment was achieved by using ClustalW (Thompson et al. 1997). Gaps were removed before further analysis was carried out. A Phi test was used to determine whether the homologous recombination event was statistically significant (Huson and Bryant 2006). Maximum-likelihood trees were constructed online using Phyml (Guindon and Gascuel 2003) (http://atgc.lirmm.fr/phyml/) and displayed using MEGA4 (Tamura et al. 2007) to determine the recombination events. Bootstrapping was employed to assess the robustness of a tree with 1,000 replications. The bootstrap values are shown below or above the branch. Phylogenetic trees were also generated using the neighbor-joining (NJ) method with maximum composite likelihood in MEGA4 (Tamura et al. 2007). The scale corresponds to the number of nucleotide substitutions per site. The Shimodaira–Hasegawa test was implemented to determine whether phylogenetic trees estimated from different regions reveal differences that are statistically significant using the Treetest program (http://aix1.uottawa.ca/~sarisbro/). Splitstree 4 was employed to find the network of mosaics and their parents (Huson and Bryant 2006).

Putative recombinant sequences were identified with the soft package of SimPlot (Lole et al. 1999) as described in a previous study (He et al. 2007). Recombination break points were identified by maximization of $\chi^2$ combined with a genetic algorithm (implementing GARD online, http://www.datamonkey.org/GARD/) (Lole et al. 1999; Kosakovsky Pond et al. 2006).

About 9,000 AIV gene sequences were also retrieved from GenBank for scanning for evidence of recombination. These AIV sequences were divided into ~100 groups before the recombination analyses were performed. In order to determine the most accurate putative parents, each mosaic was used as a query to perform Blast in GenBank.

## Results

We found that three H9N2 AIV isolates collected in different counties of the Guangxi Province at different times had similar mosaic PA segments. Among the three mosaic isolates, A/chicken/Guangxi/1/00(H9N2) also had a mosaic PB2. Moreover, we analyzed roughly 9,000 complete segments of AIVs deposited in GenBank (table 1) and at least 41 mosaics were found again (table 2 ). Homologous recombination could occur in the same subtype or between different subtypes, such as H5N1 and H5N1, H9N2 and H9N2, and H5N1 and H9N2. The recombination events were located in different genes, including *PB2*, *PB1*, *PA*, *HA*, *NA*, and *NP*. In our analysis, no mosaic was found in NS and MP segments.

### Three H9N2 AIVs Isolated from Guangxi Contain Mosaic Segment

In the year 2000, avian influenza broke out in different chicken flocks in the Guangxi Province of China. Approximately 3% of all chickens died from influenza disease in these infected chicken flocks. We isolated and purified three AIVs from three counties in March and November that were determined to be of the H9N2 subtype named A/chicken/Guangxi/1/00(H9N2), A/chicken/Guangxi/14/00(H9N2), and A/chicken/Guangxi/17/00(H9N2). The three viruses were collected from three dead birds. After each segment of the three H9N2 strains was sequenced, we found that they had very high sequence similarity in the *PA* gene (fig. 1*A*): 99.81% (2,138/2,142) between A/chicken/Guangxi/17/00 and A/chicken/Guangxi/14/00, and 99.49% (2,131/2,142) between A/chicken/Guangxi/17/00 and A/chicken/Guangxi/1/00.

**Table 2**
**AIV Strains with Evidence for Potential Recombination**

| GenBank Number | Mosaic Virus Strain (Subtype) (Reference) | Gene | Putative Parent Lineages (Subtype) | Position of Break Points | Countries |
|---|---|---|---|---|---|
| DQ073402 | A/tree sparrow/Henan/4/04(H5N1) J Virol. 79(24): 15460–15466 (2005) | *pb2* | A/tree sparrow/HN/3/04(H5N1), A/dk/Hk/293/1978(H7N2) | 1116, 1704, 1812 | China |
| DQ073399 | A/tree sparrow/Henan/1/04(H5N1) J Virol. 79(24): 15460–15466 (2005) | *pb2* | A/duck/HK/278/1978(H2N9), A/wild duck/GD/314/04(H5N1) | 636 | China |
| DQ064560 | A/chicken/jilin/53/01(H9N2) Virology. 340(1): 70–83 (2005) | *pb2* | A/ck/Yokohama/aq144/01(H9N2), A/chicken/Jilin/hk/2004(H5N1) | 1925 | China |
| EF124780 | A/chicken/Guiyang/1655/06(H5N1) Proc Natl Acad Sci USA. 103: 16936–16941 (2006) | *pa* | A/chicken/Guiyang/441/06(H5N1), A/goose Shantou/18442/05(H5N1) | 313, 595 | China |
| AY651625 | A/SCk/HK/YU100/2002(H5N1) Nature. 430(6996): 209–213 (2004) | *pa* | A/Ck/HK/31.2/2002(H5N1), A/chicken/Hubei/327/2004(H5N1) | 859,1951 | China |
| AY651512 | A/Ck/HK/37.4/2002(H5N1) Nature. 430(6996): 209–213 (2004) | *np* | A/duck/Shanghai/38/2001(H5N1), A/Ck/HK/31.4/02_(H5N1) | 799 | China |
| AF461526 | A/chicken/Tianjing/1/96(H9N2) Avian Dis. 47: 116–127 (2003) | *ha* | A/chicken/Tianjing/2/96(H9N2), A/chicken/Liaoning/1/00(H9N2) | 895 | China |
| DQ485205 | A/chicken/Guangxi/1/00(H9N2) (this study) | *pb2* | A/duck/Shanghai/13/2001(H5N1), A/chicken/Jiangsu/1/00(H9N2) | 1066 | China |
| DQ997093 | A/duck/Hubei/wg/2002(H5N1) (unpublished) | PB2 | A/swine/Anhui/ca/2004(H5N1), A/ck/Nanchang/4-301/01(H9N2) | 156, 543, 1192 | China |
| DQ997372 | A/chicken/Jilin/hq/2003(H5N1) (unpublished) | *pb2* | A/duck/Hubei/wp/2003(H5N1), A/black_duck/AUS/4045/80(H6N5) | 1887 | China |
| DQ997225 | A/chicken/Henan/wu/2004(H5N1) (unpublished) | *pb2* | A/duck/HK/278/1978(H2N9), A/chicken/Jilin/hg/2002(H5N1) | 613, 1617 | China |
| AY950280 | A/chicken/Henan/210/2004(H5N1) (unpublished) | *pb2* | A/duck/HK/278/1978(H2N9), A/chicken/Jilin/hg/2002(H5N1) | 613, 1617 | China |
| DQ997121 | A/chicken/Hubei/wj/1997(H5N1) (unpublished) | *pb2* | A/duck/HK/278/1978(H2N9), A/ck/HK/NT873.3/01-MB(H5N1) | 585 | China |
| DQ351870 | A/chicken/Hebei/718/2001(H5N1) (unpublished) | *pb2* | A/duck/Guangxi/xa/2001(H5N1), A/duck/Germany/1215/1973(H2N3) | 783, 979 | China |
| DQ997314 | A/chicken/Jilin/hh/2002(H5N1) (unpublished) | *pb2* | A/black_duck/AUS/4045/80(H6N5), A/duck/Hubei/wp/2003(H5N1) | 1638 | China |
| DQ997101 | A/chicken/Hubei/wh/1997(H5N1) (unpublished) | *pb2* | A/chicken/Jiangsu/wa/2002(H9N2), A/chicken/Hubei/wi/1997(H5N1) | 711,1323, 1695 | China |
| DQ997317 | A/chicken/Jilin/hj/2003(H5N1) (unpublished) | *pb2* | A/duck/Hongkong/d73/76(H6N1), A/chicken/Jilin/xv/2002(H5N1) | 1713 | China |
| DQ997510 | A/chicken/Beijing/ne/1999(H9N2) (unpublished) | *pb1* | A/chicken/Henan/nd/98(H9N2), A/duck/Xuzhou/07/03(H9N2) | 642, 1566 | China |
| DQ997281 | A/goose/Jilin/hb/2003(H5N1) (unpublished) | *pb1* | A/chicken/Henan/16/2004(H5N1), A/chicken/Jiangsu/wa/2002(H9N2) | 1283, 1690 | China |
| DQ997084 | A/chicken/Hubei/wf/2002(H5N1) (unpublished) | *pb1* | A/swine/Anhui/ca/2004(H5N1), A/duck/Fujian/19/2000(H5N1) | 541, 650 1234, 1963 | China |
| DQ997288 | A/chicken/Jilin/hd/2002(H5N1) (unpublished) | *pb1* | A/duck/Hubei/wp/2003(H5N1), A/swine/shandong/na/2002(H9N2) | 1182, 1843 | China |
| DQ351874 | A/chicken/Hebei/718/2001(H5N1) | *pb1* | AA/chicken/Hebei/1/2002(H7N2), A/duck/Guangxi/35/2001(H5N1) | 1638 | China |
| AY653199 | A/chicken/Jilin/9/2004(H5N1) (unpublished) | *pb1* | A/chicken/Jiangsu/cz1/2002(H5N1), A/duck/Jiangsu/nf/2003(H9N2) | 1182, 1843 | China |
| DQ997297 | A/chicken/Jilin/he/2002(H5N1) (unpublished) | *pb1* | A/chicken/Jiangsu/cz1/2002(H5N1), A/swine/shandong/na/2002(H9N2), A/duck/Hubei/W1/2004(H9N2) | 736, 1302, 1791 | China |
| DQ997449 | A/chicken/Jiangsu/nf/02(H9N2) (unpublished) | *pb1* | A/CK/Yokohama/aq45/02(H9N2), A/chicken/Jiangsu/wa/2002(H9N2) | 673 | China |
| DQ997137 | A/chicken/Hubei/wl/1997(H5N1) (unpublished) | *pa* | A/CK/Henan/210/2004_(H5N1), A/swine/Anhui/ca/2004(H5N1), one parent lineage is missing | 334, 489, 1659 | China |
| DQ351867 | A/chicken/Hebei/108/02(H5N1) (unpublished) | *pa* | A/duck/Fujian/01/2002(H5N1), A/quail/Dubai/303/2000(H9N2) | 1877 | China |
| DQ997107 | A/chicken/Hubei/wi/1997(H5N1) (unpublished) | *pa* | A/tree sparrow/Henan/2/04(H5N1), A/goose/Hujian/bb/2003 (H5N1) | 1174,1484 | China |
| DQ997274 | A/chicken/Jilin/ha/2003(H5N1) (unpublished) | *pa* | A/chicken/Jiangsu/wa/2002(H9N2), A/chicken/Jilin/9/2004(H5N1) | 1678 | China |
| DQ997280 | A/goose/Jilin/hb/2003(H5N1) (unpublished) | *pa* | A/chicken/Yamaguchi/7/2004(H5N1), A/chicken/Jiangsu/cz1/2002(H5N1), A/quail/Shantou/1461/2001(H9N2) | 674,1228, 1835 | China |
| DQ997414 | A/duck/Zhejiang/bj/2002(H5N1) (unpublished) | *pa* | A/chickem/Jilin/9/2004(H5N1), A/Gf/HK/38/2002(H5N1) | 203,1558 | China |
| DQ485223 | A/chicken/Guangxi/17/00(H9N2) (this study) | *pa* | A/chicken/Jiangsu/cz1/2002(H5N1), A/chicken/Guangdong/4/00(H9N2) | 1332 | China |

**Table 2**
**Continued**

| GenBank Number | Mosaic Virus Strain (Subtype) (Reference) | Gene | Putative Parent Lineages (Subtype) | Position of Break Points | Countries |
|---|---|---|---|---|---|
| DQ485215 | A/chicken/Guangxi/14/00(H9N2) (this study) | *pa* | A/chicken/Jiangsu/cz1/2002(H5N1), A/chicken/Guangdong/4/00(H9N2) | 1332 | China |
| DQ485207 | A/chicken/Guangxi/1/00(H9N2) (this study) | *pa* | A/chicken/Jiangsu/cz1/2002(H5N1), A/chicken/Guangdong/4/00(H9N2) | 1332 | China |
| CY014862 | A/mallard/ohio/1801/2005/(H3N8) | *pa* | A/mallard/Maryland/1131/05(H12N5), A/mallard/Maryland/124/05(H4N6) | 1343 | United States |
| DQ997303 | A/chicken/jilin/hg/02(H5N1) (unpublished) | *pa* | A/chicken/Jilin/hl/2004/(H5N1), A/chicken/Jiangsu/cz1/2002(H5N1) | 661, 1173 | China |
| DQ997321 | A/chicken/jilin/hj/03(H5N1) (unpublished) | *pa* | A/duck/Shangdong/093/04(H5N1), A/chicken/Henan/01/04(H5N1) | 687, 1173, 1749 | China |
| AF461512 | A/chicken/Gansu/1/99(H9N2) Avian Dis 47: 116–127 (2003) | *ha* | A/chicken/Osaka/aq48/97(H9N2), A/chicken/Shandong/1/98(H9N2) | 911 | China |
| AY664667 | A/CK/HongKong/NT142/03(H9N2) (unpublished) | *ha* | A/CK/HK/NT142/03(H9N2), A/CK/HongKong/WF120/03(H9N2) | 652, 1160 | China |
| DQ366322 | A/duck/Vietnam/8/05(H5N1) Arch Virol. 151(8): 1615–1624 (2006) | *ha* | A/goose/Vietnam/3/05(H5N1), the other putative parent is missing | 1086, 1400 | Vietnam |
| DQ997369 | A/chicken/Jilin/hq/2003(H5N1) (unpublished) | *np* | A/chicken/Jilin/hp/2003(H5N1), A/chicken/Henan/5/98(H9N2) | 486, 903 | China |
| AY653196 | A/chicken/Jilin/9/2004(H5N1) (unpublished) | *np* | A/chicken/Jiangsu/cz1/2002(H5N1), A/chicken/Jilin/hg/2002(H5N1) | 936, 1135 | China |
| DQ997285 | A/chicken/Jilin/hd/2002(H5N1) (unpublished) | *np* | A/chicken/Jiangsu/cz1/2002(H5N1), A/chicken/Jilin/hg/2002(H5N1) | 936, 1135 | China |
| DQ997346 | A/chicken/Jilin/hn/2003(H5N1) (unpublished) | *np* | A/chicken/jilin/xv/2002/(H5N1), A/chicken/Henan/5/98(H9N2) | 612, 1040 | China |
| DQ349117 | A/chieken/hebei/718/01(H5N1) | *na* | A/starling/England/983/79(H7N1), A/goose/Hk/3014.5/2000(H5N1) | 1090 | China |

It is very interesting to note that the three PA segments are mosaic descended from the H9N2 and H5N1 subtypes (fig. 1*B*, *C*, *E*, and *F*). When the mosaics were incorporated into a PA alignment, the evidence for recombination was found (Phi test, $P < 0.0001$). Delta distributions from Boni also gave a *P* value of 10^(−40) (Boni et al. 2007). A single break point was located at position 1332 (fig. 1*B* and *C*) using the maximized value of $\chi^2$ combined with a genetic algorithm (Lole et al. 1999; Kosakovsky Pond et al. 2006). A similarity plot constructed using all sites revealed that the PA sequence of the mosaics exhibited greater affinity with one putative parent A/chicken/Jiangsu/CZ1/02(H5N1) of the H5N1 serotype before the break point (99.32%, 1,323/1,332) (fig. 1*B*). Alternatively, the mosaics shared higher sequence similarity with A/chicken/Guangdong4/00(H9N2) of the H9N2 serotype after position 1333 (99.26%, 809/815) (fig. 1*B*). The most compelling evolutionary evidence for recombination is the occurrence of incongruent phylogenetic trees (Nelson and Holmes 2007). When maximum-likelihood phylogenetic trees before or after the break point were constructed, incorporating the human lineage, a statistically significant discrepancy between phylogenetic trees was found (Shimodaira–Hasegawa test, $P < 0.0001$) (fig. 1*E* and *F*). In phylogenetic trees, the two regions delimitated by the break point fell in different lineages with 100% bootstrap (1,000 replications); supporting the idea that the three isolates are homologous recombinants originating from the H5N1 and H9N2 lineages (fig. 1*E* and *F*). In order to see the parent strains network with the mosaics, a split tree of complete *PA* genes was constructed. The split tree suggested that the PA segment of the three H9N2 isolates descended from A/chicken/Jiangsu/CZ1/02(H5N1) and A/chicken/Guangdong4/00(H9N2) lineages (fig. 1*G*).

Before sequencing, the three isolates had been purified using a plaque-forming method to isolate a nearly clonal virus population. Additionally, several clones of each segment from the same isolate were sequenced to avoid artifacts. The mosaics shared very high sequence similarity with their putative parents before and after the break point. The high similarity is reasonable because the recombinants and their parents were contemporary (circulating between 2000 and 2002). Moreover, the major change in sequence cannot be explained by random gene mutation. Therefore, the three repeatable PA mosaics, isolated in different locations and at different times, provided robust evidence for homologous recombination as a natural phenomenon in AIVs.

In addition, PB2 of A/chicken/Guangxi/1/00(H9N2) is also a mosaic descended from the lineages of A/chicken/Guangdong/4/00(H9N2) and A/chicken/Jiangsu/cz1/02(H5N1) (fig. 2).

According to the sequence similarity (table 3), A/chicken/Guangdong4/00(H9N2) lineage could be the putative major parent of A/chicken/Guangxi/1/00(H9N2), A/chicken/Jiangsu/CZ1/02(H5N1) lineage might be the putative minor parent, and the PA of A/chicken/Guangxi/14/00(H9N2) and A/chicken/Guangxi/17/00(H9N2) seems rearranged from Guangxi1.

## Many Potential AIV Recombinants Have Been Deposited in GenBank

Homologous recombination was thought to be rare in AIVs because no mosaic AIV had been found in

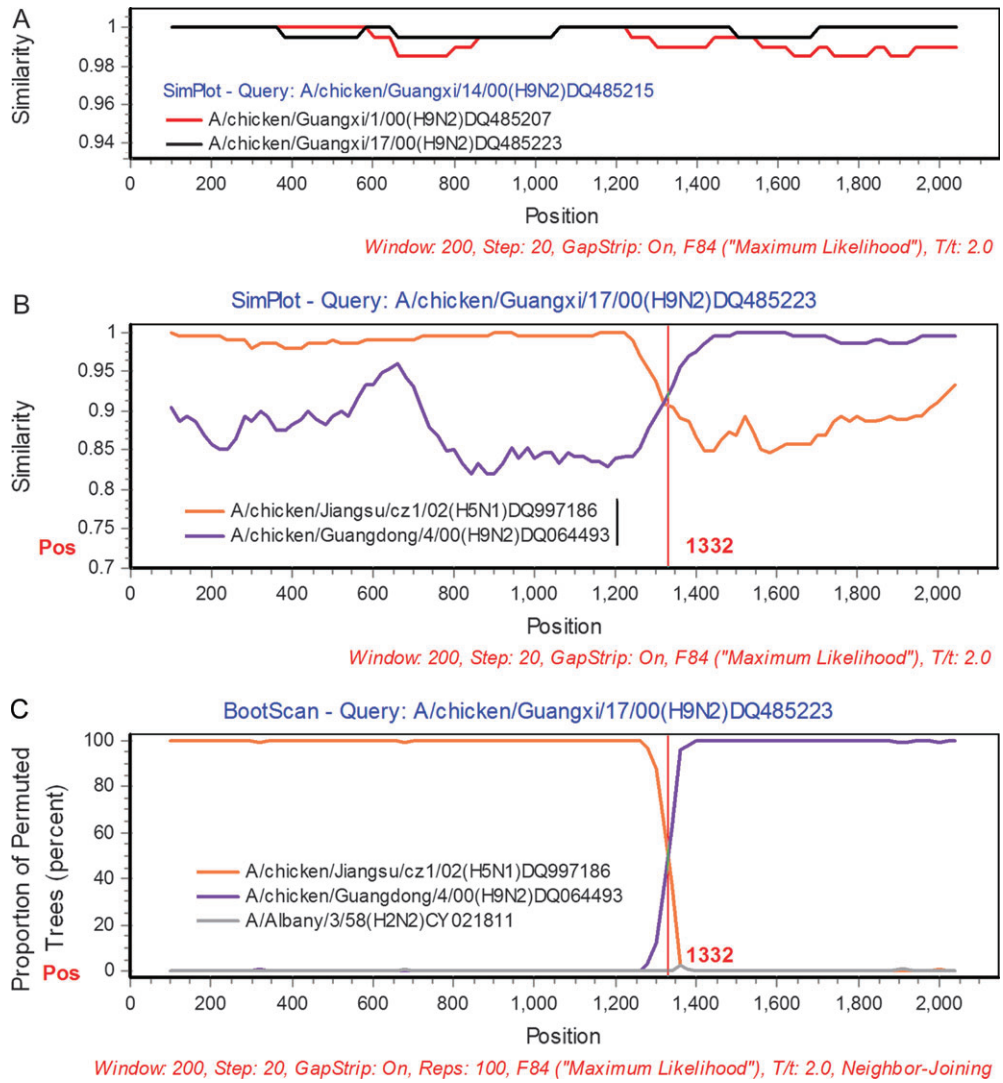


FIG. 1.—The evidence for recombination in the *PA* gene of the three H9N2 AIVs isolated from Guangxi in 2000. (*A*) A comparison of the *PA* gene from three AIV isolates: A/chicken/Guangxi/14/00(H9N2), A/chicken/Guangxi/17/00(H9N2), and A/chicken/Guangxi/1/00(H9N2). A/chicken/Guangxi/14/00(H9N2) is used as the query. The *y*-axis gives the percentage of identity within a sliding window 200 bp wide centered on the position plotted, with a step size between plots of 20 bp. (*B*) A comparison of *PA* from A/chicken/Guangxi/17/00(H9N2) and its putative parents. The red vertical line shows the recombination break point. (*C*) Bootscanning of AIV PA sequences. The *y*-axis gives the percentage of permutated trees using a sliding window. A human influenza virus isolate A/Albany/3/58(H2N2) is used as an out-group. The rest is the same as (*A*) and (*B*). (*D*), (*E*), and (*F*) are the maximum-likelihood trees of the PA segment in regions 1–2147, 1–1332, and 1333–2147, respectively. (*G*) A split tree inferred from complete PA sequence to show the evolutionary relationship between the mosaics and their parent lineages. The most accurate putative parents are shown in each parent lineage.

a large-scale sequence analysis of avian influenza isolates (Obenauer et al. 2006). To investigate this claim, we analyzed roughly 9,000 AIV sequences deposited in GenBank. Interestingly, in addition to several novel mosaic isolates in this study, we found a series of mosaics in other studies (Li et al. 2004, 2005; Kou et al. 2005; Lee et al. 2006; Smith et al. 2006) (table 2). In all, 41 mosaic segments were found among the 9,000 sequences (table 2). The whole of these segments are equal to about 1,100 (9,000/8) complete genomes of AIVs, where this means that the rate of recombination is about 3–4% on the level of the complete AIV genome deposited in GenBank.

These mosaic segments, however, are from public databases, and we cannot ascertain the detailed information of each mosaic. Therefore, these recombinants deposited in GenBank might include some artifact mosaics. The recombination rate might be overestimated here. In order to eliminate the artifacts, it is necessary to resequence these mosaic segments.

## Similar Mosaic Segments Can Be Found in Different AIV Isolates

In addition to the three repeatable PA recombinants isolated here (fig. 1), several recombinants reported by different research groups at different times shared the same recombination event and near 100% sequence similarity in the mosaic segments. These mosaic segments were PB2 of A/chicken/Henan/wu/2004(H5N1) and A/chicken/Henan/210/2004(H5N1) (fig. S34), PB2 of A/chicken/Jilin/hj/2003(H5N1) and A/chicken/Jilin/hh/2002(H5N1) (figs. S11

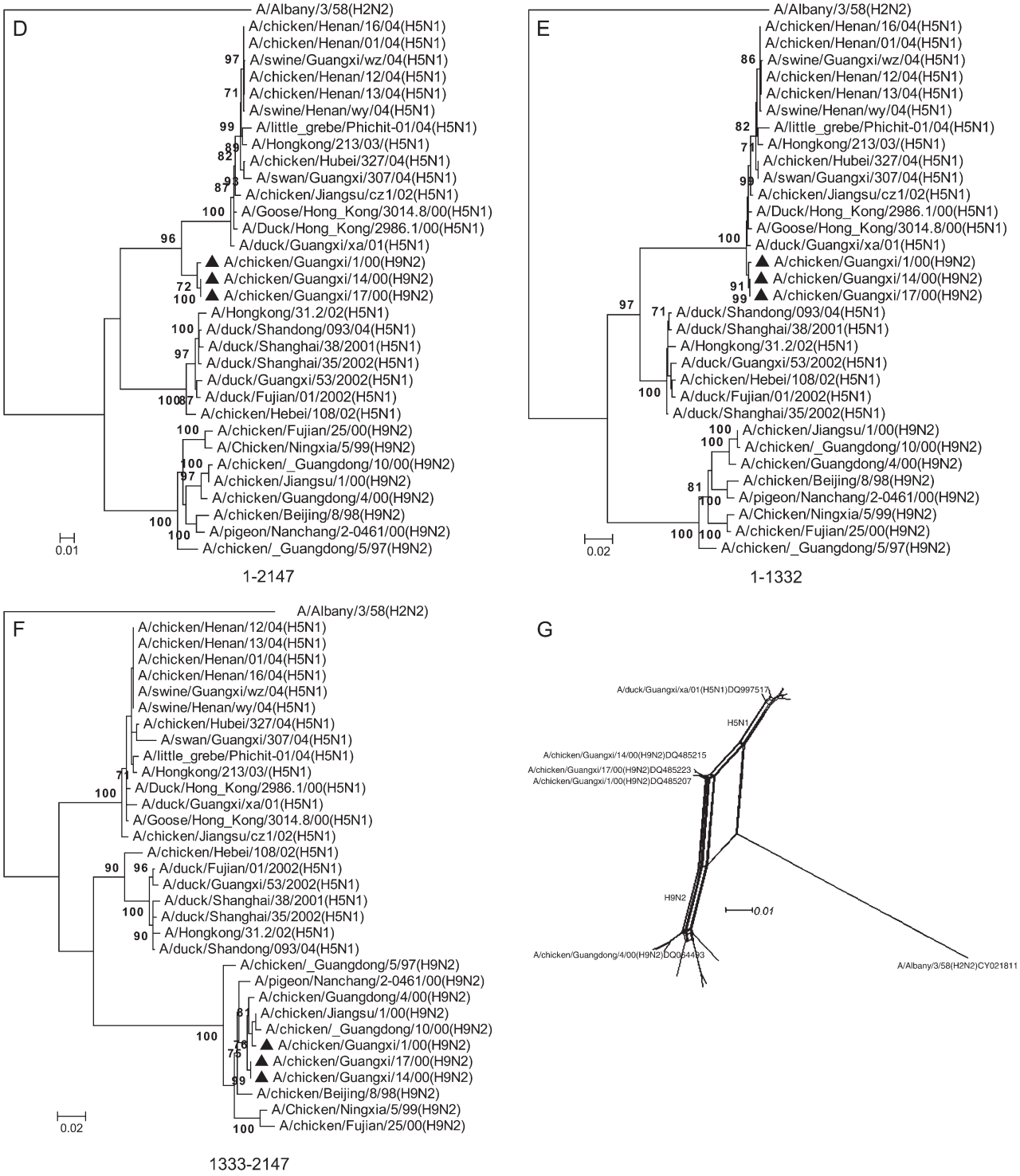


FIG. 1. (Continued).

and S12), and PB1 and NP of A/chicken/Jilin/hd/2002(H5N1) and A/chicken/Jilin/9/2004(H5N1) (figs. S16 and S27). The fact that different AIV strains containing the similar intragenic recombination segments can be isolated in different locations and at different time points by different research groups also supports the hypothesis that recombination can occur in AIVs and exist in the wild.

## Discussion

In this study, we report on three AIV recombinants isolated from the Guangxi Province of China. We also identify a series of potential recombinants deposited in GenBank. These recombinants collectively demonstrate that homologous recombination can play a role in shaping genetic diversity of the virus.

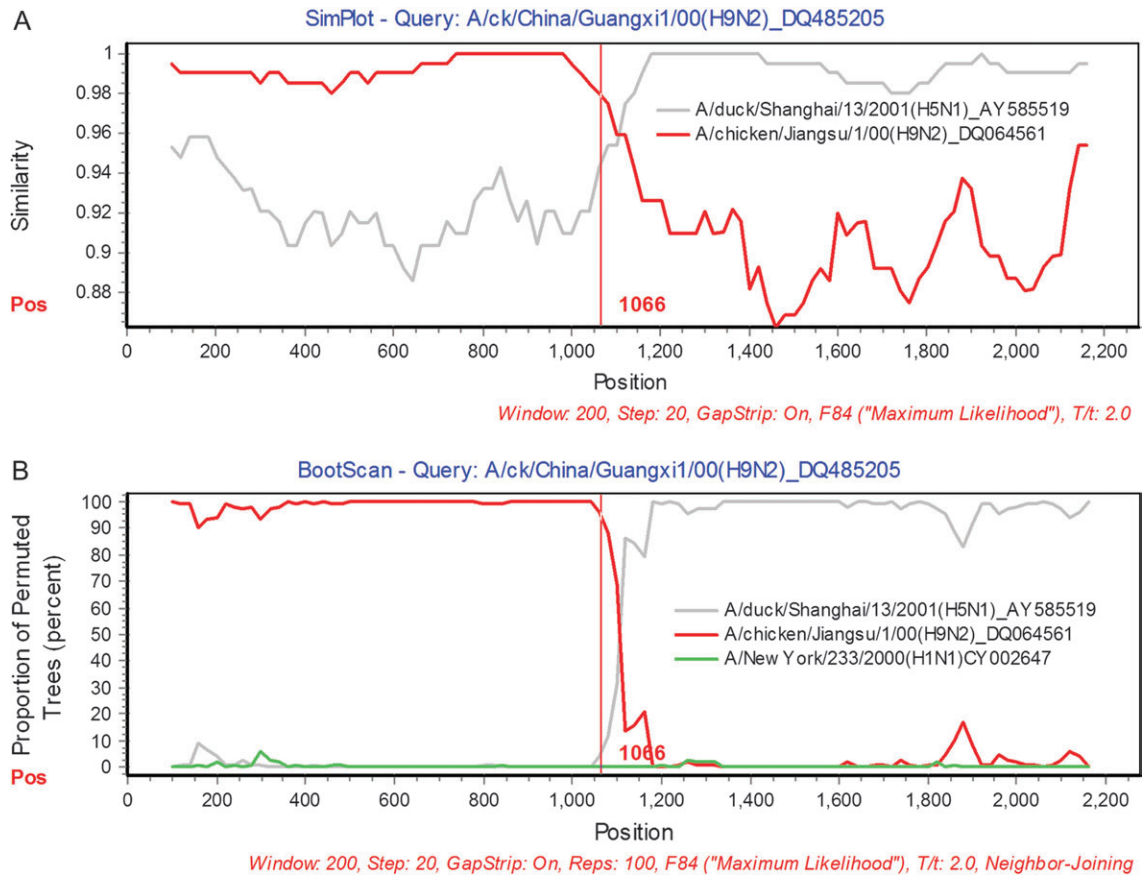


FIG. 2.—The evidence for recombination in PB2 segment of A/chicken/Guangxi/1/00(H9N2). Results from SimPlot analysis of A/chicken/Guangxi/1/00(H9N2) are shown. The *y*-axis gives the percentage of identity within a sliding window 200 bp wide centered on the position plotted, with a step size between plots of 20 bp. The analysis was carried out using A/chicken/Guangxi/1/00(H9N2) as a query. The two putative parent lineages are shown with different colors. The red vertical lines show the recombination break points with the maximization of $\chi^2$. (B) Bootscanning of AIV PB2 sequences. The *y*-axis gives the percentage of permutated trees using a sliding window 200 bp wide centered on the position plotted, with a step size between plots of 20 bp. The rest is the same as (*A*). (*C*) and (*D*) are the ML phylogenetic trees for the recombinant regions from 1 to 1066 and 1067 to 2277, respectively. A human influenza virus isolate A/New York/233/2000(H1N1) is used as the out-group. (*F*) A split tree inferred from complete PB2 sequence to show the PB2 evolution relation between A/chicken/Guangxi/1/00(H9N2) and their parent lineages. The most accurate putative parents are shown in each parent lineage.

The three AIVs isolated in Guangxi contained mosaic PAs sharing very high sequence similarity to each other (fig. 1*A*). The same recombination event and phylogenetic evidence (fig. 1*B*–*G*) reveal that these PAs descended from the same H9N2/H5N1 recombinant ancestor. Guangxi1 was collected on 20 March, whereas Guangxi14 and Guangxi17 were collected on 13 November and 20 November, respectively. Therefore, the PAs of Guangxi14 and Guangxi17 might be descended from the Guangxi1 lineage, which is supported by the phylogenetic relationship of the three mosaics (fig. 1*D*). Similar observations can also be found in PB2 of A/chicken/Jilin/hj/2003(H5N1) and A/chicken/Jilin/hh/2002(H5N1) (fig. S11) and PB2 of A/chicken/Henan/wu/2004(H5N1) and A/chicken/Henan/210/2004(H5N1) (fig. S34). The fact that different isolates, originating from the same mosaic ancestor, can be found in different locations and at different time points provides robust evidence that intragenic recombination between different AIV serotypes results in novel viruses. Moreover, these viruses containing mosaic segment(s) are stable and can circulate in the field. Therefore, intragenic recombination does act as a force in the evolution of AIVs.

The recombination, especially between different subtypes, will speed up evolution of the virus. After the AIV H9N2 subtype was first isolated in China in 1994, it became responsible for the 93.89% of avian influenza outbreaks in the country from 1996 to 2000 (Guo et al. 2000). The H5N1 virus is now endemic in poultry in Asia (Li et al. 2004), and the outbreaks occur from time to time in China. This indicates that the two subtypes are circulating at the same time in China, which provides the means for recombination to occur between them. We found that three mosaics descended from H5N1 and H9N2 in the Guangxi Province (figs. 1 and 2). Similar H9N2/H5N1 recombinants can be also found in GenBank (table 2). Both of the subtypes can infect humans, although infection with H9N2 is thought to be nonfatal (Peiris et al. 1999). It is therefore important to evaluate the effects on H9N2 virulence as a result of recombination between subtypes.

Interestingly, almost all recombinants were isolated in China (table 2). It seems that the AIV has become entrenched in its ecological niche where homologous recombination can occur, highlighting a potential long-term

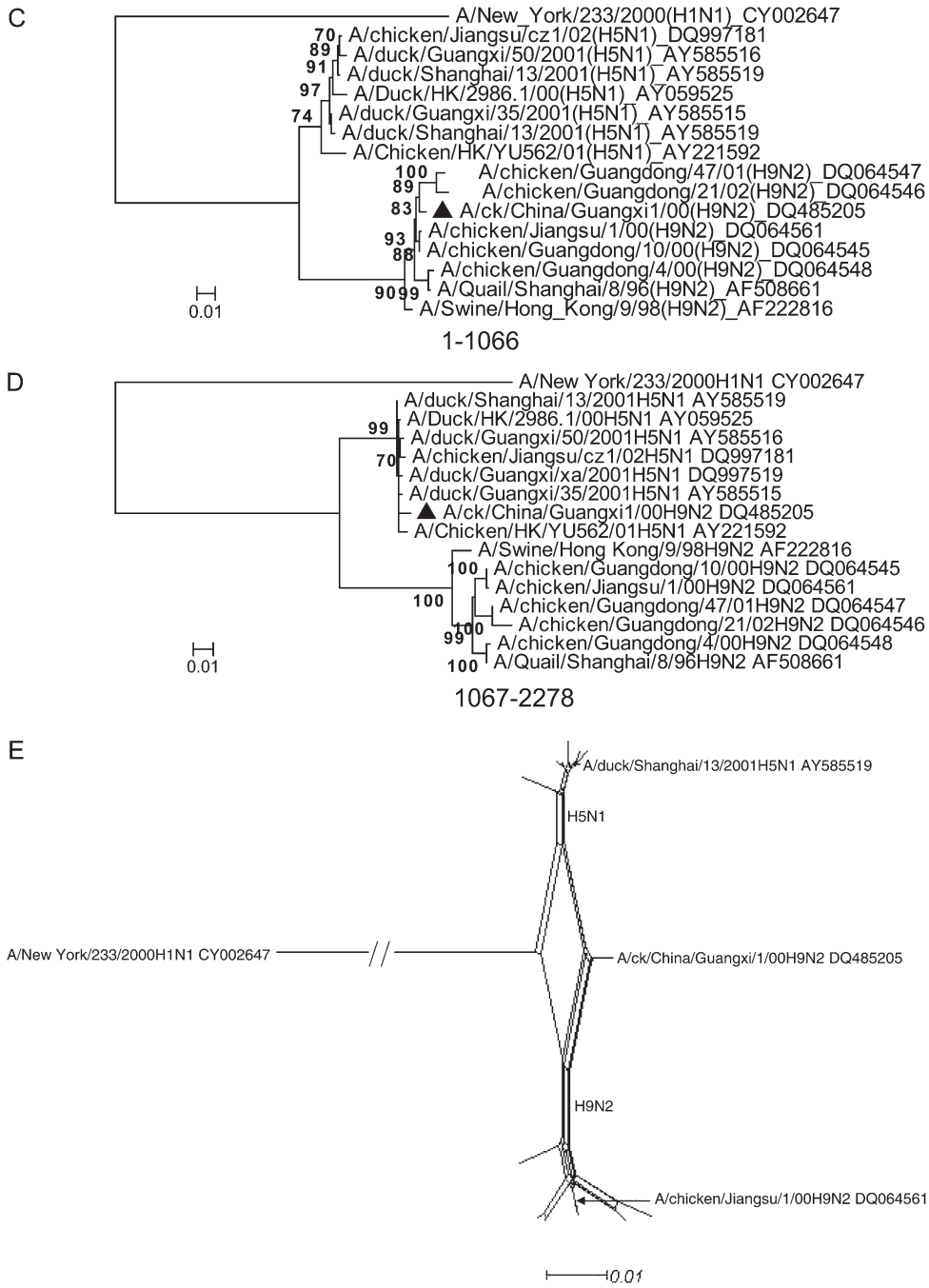


FIG. 2. (Continued).

pandemic threat to humans. At present, the mechanism of transmission of these viruses from poultry to humans is poorly understood (Li et al. 2004). In evolutionary terms, and from the perspective of the pathogen, the host species barrier for infection can be thought of as a fitness valley lying between two distinct fitness peaks, representing donor and recipient hosts, respectively (Kuiken et al. 2006). The mutation in polymerase genes was thought to be responsible for the AIV transmission to human populations in the 1918 influenza pandemic (Taubenberger et al. 2005). Reassortment and recombination processes will allow some viruses to acquire many of the key adaptive mutations in a single step and thus make a major leap in fitness, which might result in a change of host tropism (Kuiken et al. 2006). A re-assortment of gene segments between pig and human influenza has resulted in a novel virus to which humans were immunologically naive (Russell and Webster 2005). Similarly, the intragenic recombination between AIV and human influenza virus is also capable of resulting in a novel virus, adaptable to humans. In fact, intragenic recombination has been found to occur between polymerase members in this analysis, which are known to be involved

**Table 3**
**Comparison of the Putative Recombinants Isolated from Guangxi and Their Putative Parents**

| | Similarity | | | | | |
|---|---|---|---|---|---|---|
| Segment | GX1/ GD4 | GX1/ JS | GX14/ GD4 | GX14/ JS | GX17/ GD4 | GX17/ JS |
| HA | 99.31 | 57.6 | 98.51 | 57.7 | 96.84 | 56.13 |
| PA | 92.92 | 95.48 | 92.88 | 95.43 | 92.97 | 95.52 |
| PB1 | 99.18 | 90.56 | 89.31 | 91.25 | 89.27 | 91.12 |
| PB2 | 94.8 | 95.96 | 88.97 | 93 | 85.66 | 86.6 |
| NP | 98.4 | 90.12 | 88.83 | 92.09 | 89.22 | 92.04 |
| NA | 98.4 | 49.14 | 88.3 | 50.96 | 89.2 | 51.1 |
| M | 99.61 | 92.3 | 98.44 | 91.82 | 99.61 | 92.31 |
| NS | 98.2 | 90.63 | 95.17 | 90.63 | 94.83 | 90.29 |

NOTE.—GX1, A/chicken/Guangxi/1/00(H9N2); GX17, A/chicken/Guangxi/17/00(H9N2); GX14, A/chicken/Guangxi/14/00; JS, A/chicken/Jiangsu/CZ1/02(H5N1); GD, A/chicken/Guangdong4/00(H9N2).

**Table 4**
**The Nucleotide Sequences of the Specific Primer Pairs Used in the Study**

| Gene | Sequence of Primers |
|---|---|
| *pb2* | F: 5′-AAAAGCAGGTCAATTATATTC-3′; R: 5′-AAGGTCGTTTTTAAACTATTCA-3′ |
| *pb1* | F: 5′-AAAAGCAGGCAAACCATTTGA-3′; R: 5′-TTTTCATGAAGGACAAGCTAA-3′ |
| *pa* | F: 5′-AGCAAAAGCAGGTACTGAT-3′; R: 5′-AGTAGAAACAAGGTACTTTT-3′ |
| *ha* | F: 5′-AGCAAAAGCAGGGGAATTTCAC-3′; R: 5′-AGTAGAAAACAAGGGTGTTTTTGC-3′ |
| *np* | F: 5′-GCAGGTAGATAATCACTCACTG-3′; R: 5′-AGTAGAAACAAGGGTATTTTT-3′ |
| *na* | F: 5′-AGCAAAAGCAGGAGTAAAAATG-3′; R: 5′-CAAGGAGTTTTTTTTTAAAATTGC-3′ |
| *m* | F: 5′-AGCAAAAGCAGGTAGATGTTTAAAG-3′; R: 5′-AGTAGAAACAAGGTAGTTTTTTAC-3′ |
| *ns* | F: 5′-AAAGCAAGGGTGACAAAGACAT-3′; R: 5′-TAGAAACAAGGGTGTTTTTTATCA-3′ |

NOTE.—F, forward; R, reverse.

in many aspects of viral replication and to interact with host factors, thus having a role in host specificity (Taubenberger et al. 2005). PB1 is associated with avian-to-human transmission of the *PB1* gene of influenza A viruses in the 1957 and 1968 pandemics (Kawaoka et al. 1989). Although we did not find the evidence that recombination occurred between human and avian influenza viruses, an AIV strain isolated from swine seems to be able to act as a putative parent (table 2). We have also found that the recombination can occur between human and swine influenza viruses (He et al. 2008). Therefore, proper attention should be given to the potential of homologous recombination events triggering pandemics by altering gene structure or function and/or permitting the highly virulent virus to switch hosts, from birds to humans.

Recently, homologous recombination has also been reported in several negative-strand RNA viruses, such as ambisense arenaviruses (Charrel et al. 2001; Archer and Rico-Hesse 2002), the human respiratory syncytial virus (Spann et al. 2003), and the Newcastle disease virus (Han et al. 2008; Qin et al. 2008). Because there is evidence to show that influenza viruses undergo various forms of nonhomologous recombination (Khatchikian et al. 1989; Orlich et al. 1994), there should be opportunities for intragenic recombination in the evolution of AIVs. However, the molecular mechanism underlying this recombination is unclear in AIVs. For positive RNA viruses, template switching has been reported as the mechanism of intragenic recombination and as the main mechanism in many cases (Lai 1992). For AIVs, however, it is known that the synthesis of RNA strands of the genome is particle associated, and free double-stranded or minus-strand RNA has never been found in infected cells. This seems to reduce the probability of both homologous recombination and template switching occurring in the replication process of the virus. In fact, we also found that the rate of recombination was not as high as that in a positive-sense RNA virus, for example, the foot-and-mouth disease virus in which 10–20% of viral genomes undergo recombination during a single replication cycle (Alejska et al. 2001). In studies on intertypic poliovirus, recombination occurred in RNA regions where RNA could potentially form a secondary structure and play an activator role in recombination (Nagy et al. 1999). This allows the two parental RNAs of different origins to form a complex and thereby force recombination to occur (Romanova et al. 1986; Tolskaya et al. 1987). Therefore, analysis of the characteristics of break points can provide clues about the recombination mechanism in AIVs. When the recombination break point sequences were analyzed, a potential secondary structure of the region around the break point was also found in some mosaics (data not shown), which could provide a molecular basis for the recombination to occur.

In conclusion, this study provides evidence demonstrating that recombination can occur in AIVs and plays a role in the evolution of the virus. It will be important to evaluate its influence on virulence and host tropism and to study the mechanism of AIV recombination.

## Supplementary Material

Supplementary figures are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

## Literature Cited

Alejska M, Kurzyniska-Kokorniak A, Broda M, Kierzek R, Figlerowicz M. 2001. How RNA viruses exchange their genetic material. Acta Biochim Pol. 48:391–407.

Anderson JP, Rodrigo AG, Learn GH, Madan A, Delahunty C, Coon M, Girard M, Osmanov S, Hood L, Mullins JI. 2000. Testing the hypothesis of a recombinant origin of human immunodeficiency virus type 1 subtype E. J Virol. 74: 10752–10765.

Archer AM, Rico-Hesse R. 2002. High genetic divergence and recombination in Arenaviruses from the Americas. Virology. 304:274–281.

Becher P, Orlich M, Thiel HJ. 2001. RNA recombination between persisting pestivirus and a vaccine strain: generation of cytopathogenic virus and induction of lethal disease. J Virol. 75:6256–6264.

Boni MF, Posada D, Feldman MW. 2007. An exact nonparametric method for inferring mosaic structure in sequence triplets. Genetics. 176:1035–1047.
Boni MF, Zhou Y, Taubenberger JK, Holmes EC. 2008. Homologous recombination is very rare or absent in human influenza A virus. J Virol. 82:4807–4811.
Chang S, Zhang J, Liao X, et al. (12 co-authors). 2007. Influenza Virus Database (IVDB): an integrated information resource and analysis platform for influenza virus research. Nucleic Acids Res. 35:D376–D380.
Charrel RN, de Lamballerie X, Fulhorst CF. 2001. The Whitewater Arroyo virus: natural evidence for genetic recombination among Tacaribe serocomplex viruses (family Arenaviridae). Virology. 283:161–166.
Gibbs MJ, Armstrong JS, Gibbs AJ. 2001. Recombination in the hemagglutinin gene of the 1918 "Spanish flu". Science. 293: 1842–1845.
Gorman OT, Bean WJ, Kawaoka Y, Webster RG. 1990. Evolution of the nucleoprotein gene of influenza A virus. J Virol. 64:1487–1497.
Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 52:696–704.
Guo YJ, Krauss S, Senne DA, Mo IP, Lo KS, Xiong XP, Norwood M, Shortridge KF, Webster RG, Guan Y. 2000. Characterization of the pathogenicity of members of the newly established H9N2 influenza virus lineages in Asia. Virology. 267:279–288.
Han GZ, He CQ, Ding NZ, Ma LY. 2008. Identification of a natural multi-recombinant of Newcastle disease virus. Virology. 371:54–60.
Hatta M, Gao P, Halfmann P, Kawaoka Y. 2001. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. Science. 293:1840–1842.
He CQ, Ding NZ, Fan W, Wu YH, Li JP, Li YL. 2007. Identification of chicken anemia virus putative intergenotype recombinants. Virology. 366:1–7.
He CQ, Han GZ, Wang D, Liu W, Li GR, Liu XP, Ding NZ. 2008. Homologous recombination evidence in human and swine influenza A viruses. Virology. 380:12–20.
Horimoto T, Kawaoka Y. 2005. Influenza: lessons from past pandemics, warnings from current incidents. Nat Rev Microbiol. 3:591–600.
Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 23:254–267.
Kawaoka Y, Krauss S, Webster RG. 1989. Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. J Virol. 63:4603–4608.
Khatchikian D, Orlich M, Rott R. 1989. Increased viral pathogenicity after insertion of a 28S ribosomal RNA sequence into the haemagglutinin gene of an influenza virus. Nature. 340:156–157.
Kirkegaard K, Baltimore D. 1986. The mechanism of RNA recombination in poliovirus. Cell. 47:433–443.
Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. Mol Biol Evol. 23:1891–1901.
Kou Z, Lei FM, Yu J, et al. (11 co-authors). 2005. New genotype of avian influenza H5N1 viruses isolated from tree sparrows in China. J Virol. 79:15460–15466.
Kuiken T, Holmes EC, McCauley J, Rimmelzwaan GF, Williams CS, Grenfell BT. 2006. Host species barriers to influenza virus infections. Science. 312:394–397.
Lai MM. 1992. RNA recombination in animal and plant viruses. Microbiol Rev. 56:61–79.
Lee MS, Chang PC, Shien JH, Cheng MC, Chen CL, Shieh HK. 2006. Genetic and pathogenic characterization of H6N1 avian influenza viruses isolated in Taiwan between 1972 and 2005. Avian Dis. 50:561–571.
Li C, Yu K, Tian G, Yu D, Liu L, Jing B, Ping J, Chen H. 2005. Evolution of H9N2 influenza viruses from domestic poultry in Mainland China. Virology. 340:70–83.
Li KS, Guan Y, Wang J, et al. (21 co-authors). 2004. Genesis of a highly pathogenic and potentially pandemic H5N1 influenza virus in eastern Asia. Nature. 430:209–213.
Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol. 73:152–160.
Nagy PD, Ogiela C, Bujarski JJ. 1999. Mapping sequences active in homologous RNA recombination in brome mosaic virus: prediction of recombination hot spots. Virology. 254:92–104.
Nagy PD, Simon AE. 1997. New insights into the mechanisms of RNA recombination. Virology. 235:1–9.
Nelson MI, Holmes EC. 2007. The evolution of epidemic influenza. Nat Rev Genet. 8:196–205.
Obenauer JC, Denson J, Mehta PK, et al. (17 co-authors). 2006. Large-scale sequence analysis of avian influenza isolates. Science. 311:1576–1580.
Orlich M, Gottwald H, Rott R. 1994. Nonhomologous recombination between the hemagglutinin gene and the nucleoprotein gene of an influenza virus. Virology. 204: 462–465.
Peiris M, Yuen KY, Leung CW, Chan KH, Ip PL, Lai RW, Orr WK, Shortridge KF. 1999. Human infection with influenza H9N2. Lancet. 354:916–917.
Pita JS, Fondong VN, Sangare A, Otim-Nape GW, Ogwal S, Fauquet CM. 2001. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. J Gen Virol. 82:655–665.
Qin Z, Sun L, Ma B, Cui Z, Zhu Y, Kitamura Y, Liu W. 2008. F gene recombination between genotype II and VII Newcastle disease virus. Virus Res. 131:299–303.
Romanova LI, Blinov VM, Tolskaya EA, Viktorova EG, Kolesnikova MS, Guseva EA, Agol VI. 1986. The primary structure of crossover regions of intertypic poliovirus recombinants: a model of recombination between RNA genomes. Virology. 155:202–213.
Russell CJ, Webster RG. 2005. The genesis of a pandemic influenza virus. Cell. 123:368–371.
Smith GJ, Fan XH, Wang J, et al. (13 co-authors). 2006. Emergence and predominance of an H5N1 influenza variant in China. Proc Natl Acad Sci U S A. 103:16936–16941.
Spann KM, Collins PL, Teng MN. 2003. Genetic recombination during coinfection of two mutants of human respiratory syncytial virus. J Virol. 77:11201–11211.
Steinhauer DA, Skehel JJ. 2002. Genetics of influenza viruses. Annu Rev Genet. 36:305–332.
Tamura KJ, Dudley, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.
Taubenberger JK, Reid AH, Lourens RM, Wang R, Jin G, Fanning TG. 2005. Characterization of the 1918 influenza virus polymerase genes. Nature. 437:889–893.
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25:4876–4882.
Tolskaya EA, Romanova LI, Blinov VM, Viktorova EG, Sinyakov AN, Kolesnikova MS, Agol VI. 1987. Studies on the recombination between RNA genomes of poliovirus: the primary structure and nonrandom distribution of crossover

regions in the genomes of intertypic poliovirus recombinants. Virology. 161:54–61.
Worobey M, Rambaut A, Holmes EC. 1999. Widespread intra-serotype recombination in natural populations of dengue virus. Proc Natl Acad Sci U S A. 96:7352–7357.
Worobey M, Rambaut A, Pybus OG, Robertson DL. 2002. Questioning the evidence for genetic recombination in the 1918 "Spanish flu" virus. Science. 296:211.discussion 211
Xie Z, Fadl AA, Girshick T, Khan MI. 1997. Amplification of avian reovirus RNA using the reverse transcriptase-polymerase chain reaction. Avian Dis. 41:654–660.

Robin Bush, Associate Editor

Accepted October 12, 2008