

# SCIENTIFIC REPORTS



OPEN

## Emergence of BA9 genotype of human respiratory syncytial virus subgroup B in China from 2006 to 2014

Jinhua Song<sup>1</sup>, Huiling Wang<sup>1</sup>, Jing Shi<sup>1,2</sup>, Aili Cui<sup>1</sup>, Yanzhi Huang<sup>3</sup>, Liwei Sun<sup>3</sup>, Xingyu Xiang<sup>4</sup>, Chaofeng Ma<sup>5</sup>, Pengbo Yu<sup>6</sup>, Zifeng Yang<sup>7</sup>, Qi Li<sup>8</sup>, Teresa I. Ng<sup>9</sup>, Yan Zhang<sup>1</sup>, Rongbo Zhang<sup>10</sup> & Wenbo Xu<sup>1,10</sup>

A study was conducted to investigate the circulation of HRSV subgroup B (HRSVB) in China in recent years. HRSVB sequences from 365 samples collected in 1991, 2004 and 2008–2014 in China, together with 332 Chinese HRSVB sequences obtained from GenBank were analyzed to determine the geographic and yearly distribution of HRSVB. Phylogenetic analysis revealed these HRSVB sequences clustered into 4 genotypes with different frequencies: BA (83%), CB1 (11%), SAB (3.0%) and GB3 (0.7%). Between 2005 and 2013, there was a co-circulation of BA and non-BA genotypes in China. Genotypes BA9 and BA10 were two of the main BA genotypes detected in this study. Genotype BA9 was first detected in China in 2006 and became the predominant HRSVB genotype circulating in China from 2008 to 2014. Three different lineages were detected for both genotypes BA9 and BA10. Time to the most recent common ancestor for genotypes BA9 and BA10 was estimated for years 1997 and 1996, respectively. Results of this study not only contribute to the understanding of the circulation pattern, but also the phylogenetic pattern and evolution of HRSVB in China from 1991 to 2014.

Human respiratory syncytial virus (HRSV) is the major cause of acute lower respiratory tract infection (ALRTI) worldwide in infants and young children (<5 years of age), as well as in the elderly and patients who are immunocompromised<sup>1–3</sup>. In China, among 28,369 patients with ALRTI from 81 sentinel hospitals in 22 provinces, HRSV was the most frequently detected virus (9.9%) and was also the most common etiology of ALRTI in children <2 years of age in China<sup>4–6</sup>.

HRSV is a member of the *Paramyxoviridae* family and the *Pneumovirinae* subfamily. The HRSV virion consists of a non-segmented, single-strand negative RNA genome packaged in a lipid envelope. The genome of HRSV is about 15.2 kb in length and encodes 11 proteins: NS1, NS2, N, P, M, SH, G, F, M2-1, M2-2, and L. The F and G proteins are the most important viral transmembrane surface glycoproteins. The F protein is highly conserved and the G protein is highly variable. According to the reactivity with monoclonal antibodies against surface glycoproteins, HRSV strains have been classified as subgroup A (HRSVA) or subgroup B (HRSVB)<sup>7</sup>. To date, based on the sequence of the second hypervariable region (HVR2) of the G protein, HRSVA has been divided into 14

<sup>1</sup>WHO WPRO Regional Reference Measles/Rubella Laboratory and Key Laboratory of Medical Virology, Ministry of Health, National Institute for Viral Disease Control and Prevention, China Center for Disease Control and Prevention, Beijing, People's Republic of China. <sup>2</sup>Lu Juan Community Health Center of Daxing region, Beijing, People's Republic of China. <sup>3</sup>Jilin Children's Medical Center, Children's Hospital of Changchun, Changchun, People's Republic of China. <sup>4</sup>Hunan Provincial Centers for Disease Control and Prevention, Changsha, People's Republic of China. <sup>5</sup>Xian Center for Disease Control and Prevention, Xian, People's Republic of China. <sup>6</sup>Shaanxi Provincial Centers for Disease Control and Prevention, Xian, People's Republic of China. <sup>7</sup>State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, First Affiliated Hospital of Guangzhou Medical University, Guangzhou, People's Republic of China. <sup>8</sup>Hebei Provincial Centers for Disease Control and Prevention, Shijiazhuang, People's Republic of China. <sup>9</sup>AbbVie, Inc, North Chicago, IL, USA. <sup>10</sup>Medical College, Anhui University of Science & Technology, Huainan, People's Republic of China. Correspondence and requests for materials should be addressed to Y.Z. (email: [zhangyan9876543@163.com](mailto:zhangyan9876543@163.com)) or R.Z. (email: [lory456@126.com](mailto:lory456@126.com)) or W.X. (email: [wenbo\\_xu1@aliyun.com](mailto:wenbo_xu1@aliyun.com))

Region	Dongbei	Huabei		Xibei		Xinan		Zhongnan			Huadong			Total
Province/City	Jilin	Beijing	Hebei	Shaanxi	Gansu	Sichuan	Chongqing	Hong Kong	Guangdong	Hunan	Zhejiang	Shandong	Shanghai	Total
1991	2(2)	—	—	—	—	—	—	—	—	—	—	—	—	2(2)
2000	—	—	—	—	—	—	—	3(0)	—	—	—	—	—	3(0)
2001	—	—	—	—	—	—	—	3(0)	—	—	—	—	—	3(0)
2004	—	4(2)	—	—	—	—	—	4(0)	—	—	—	—	—	8(2)
2005	—	5(0)	—	—	—	—	—	6(0)	—	—	—	—	—	11(0)
2006	—	—	—	—	—	—	4(0)	3(0)	—	—	—	—	—	7(0)
2007	—	3(0)	—	—	—	—	1(0)	1(0)	—	—	—	—	—	5(0)
2008	—	16(11)	—	—	28(0)	—	15(0)	2(0)	—	—	—	—	—	61(11)
2009	—	144(107)	—	—	1(1)	—	11(0)	5(0)	2(2)	—	—	—	16(0)	179(110)
2010	—	48(26)	1(1)	2(2)	2(2)	8(0)	5(0)	9(0)	9(9)	—	—	—	27(0)	111(40)
2011	—	7(0)	—	—	—	6(0)	6(0)	—	1(1)	—	—	—	1(0)	21(1)
2012	2(2)	5(0)	—	—	—	4(0)	2(0)	—	2(0)	11(11)	—	—	2(0)	28(13)
2013	34(34)	—	3(3)	30(30)	—	22(0)	—	—	5(0)	13(13)	28(0)	—	—	135(80)
2014	62(62)	15(15)	1(1)	8(8)	—	11(0)	—	—	—	24(18)	—	2(2)	—	123(106)
Total	100(100)	247(161)	5(5)	40(40)	31(3)	51(0)	44(0)	36(0)	19(12)	48(42)	28(0)	2(2)	46(0)	697(365)

**Table 1.** Distribution of HRSVB samples collected in China from 1991 to 2014. Numbers shown are the sum of sequences collected in this study and those downloaded from GenBank; sequences collected in this study are shown within parentheses. The 6 representatives regions of China in our study were Dongbei (Jilin), Huabei (Beijing and Hebei), Xibei (Shaanxi and Gansu), Xinan (Sichuan and Chongqing), Zhongnan (Guangdong, Hunan and Hong Kong SAR), and Huadong (Shandong, Shanghai and Zhejiang).

genotypes, GA1-GA7<sup>8</sup>, SAA1<sup>9</sup>, CB-A<sup>10</sup>, NA1-4<sup>11</sup> and ON1<sup>12</sup>, whereas HRSVB has been divided into 24 genotypes, GB1-GB4<sup>8</sup>, SAB1-SAB4<sup>13</sup>, URU1-2<sup>14</sup>, BA1-12<sup>10,15,16</sup>, CB1<sup>11</sup>, GB5<sup>17</sup>, and CBB<sup>10</sup>.

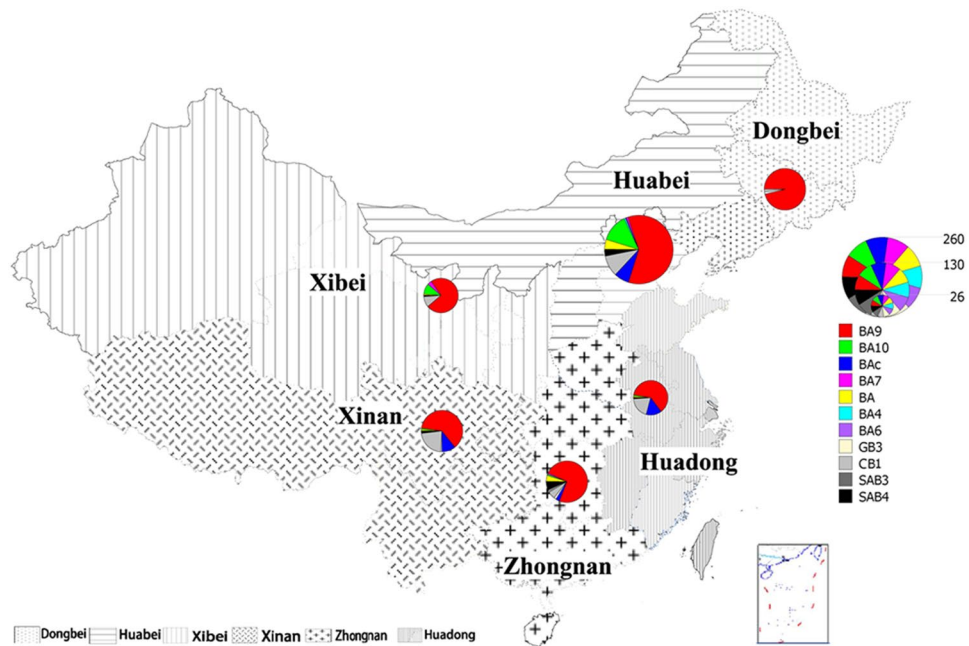
In a previous study, we found that HRSVB predominated between 2008 and 2010 in China while HRSVA predominated in 2010 and 2012, and both subgroups co-circulated in China till March 2014<sup>18</sup>. We also reported that the ON1 genotype became the predominant genotype of HRSVA circulating in China between 2013 and 2015. In this report, we describe the molecular epidemiology, circulation pattern, and evolution of different HRSVB genotypes in 6 representative regions of China from 1991 to 2014. One of the HRSVB genotypes circulating worldwide in this period of time was the BA genotype, which was first detected with a 60-nucleotide duplication in the C-terminal of the G gene in 1998. Since then, the BA genotype has spread worldwide, and 12 genotypes of BA have been identified<sup>10,15,16,19,20</sup>. While the BA genotype has been reported circulating in different local regions in China including Gansu<sup>21</sup>, Beijing<sup>11</sup>, Chongqing<sup>17,22</sup>, Shanghai<sup>23</sup>, Zhejiang<sup>24</sup> and Sichuan<sup>25</sup> during the epidemic seasons from 2006 to 2013, the epidemiology of all HRSVB genotypes in China during this period time has not been studied extensively. In this study, the circulation and evolution of Chinese HRSVB genotypes were analyzed and this study represents one of the biggest and most extensive studies to investigate the circulation of HRSVB in different parts of China in the past decade.

## Results

**Samples and sequences information.** 4115 samples were collected from patients diagnosed with respiratory infection from 5 representative regions of China from 2008 to 2014. A total of 656 out of these 4115 (15%) samples were identified as positive for HRSV. Of these, 295 samples were identified as HRSVA and 360 were identified as HRSVB, and 1 sample was identified to be positive for both HRSVA and HRSVB. Analysis of the HRSVA sequences has been reported previously<sup>18</sup>. In this study, 361 HRSVB sequences, together with sequences from 2 HRSVB samples collected from Jilin in 1991, as well as sequences from 2 HRSVB samples collected from Beijing in 2004, were analyzed.

In addition, 332 Chinese HRSVB sequences were downloaded from GenBank. Therefore, a total of 697 sequences from 13 provinces/cities of 6 geographically distinct regions of China were analyzed to determine the molecular epidemiology of HRSVB in China. Among these 697 sequences, 100 of them were obtained from the Dongbei region, 252 were obtained from the Huabei region, 76 were collected from the Huadong region, 71 were collected from the Xibei region, 95 were obtained from the Xinan region, and 103 were collected from the Zhongnan region (Table 1 and Fig. 1).

**Clinical characteristic of HRSVB infection.** Clinical information was available for 317 out of the 365 patients from whom the HRSVB samples were collected in this study. Most (92%, n = 292) of the 317 samples were collected from inpatients and 8% (n = 25) were from outpatients. The ages of the HRSVB-infected patients ranged from 1 day to 68 years old; the majority (97%) of the patients were infants or children younger than 5 years of age, including 127 (40%) females and 190 (60%) males. For the clinical diagnoses, bronchitis with pneumonia was the leading diagnosis with 158 (50%) cases, 104 (33%) were diagnosed with pneumonia, 7% were diagnosed with upper respiratory tract infection, and 2% were diagnosed with other diseases (1 case with appendicitis, 1 case with enteritis and 3 cases with unknown diagnoses). Complications occurred mainly in children under 2 years of



**Figure 1.** Geographic distribution of HRSVB samples in 6 representative regions of China. The region and the number of sequences are colored according to the legend. The map was constructed using Mapinfo Professional software, Version 11.0 (<http://www.pitneybowes.com/us/location-intelligence/geographic-information-systems/mapinfo-pro.html>).

age, with 12% (39/317) having one or more complications. The most common complications were heart disease 38% (15/39) and respiratory failure 15% (6/39). Other complications included thrush, diarrhea, asthma, biliary atresia and atelectasis.

**Phylogenetic analysis of Chinese HRSVB sequences.** All of the HRSVB sequences were classified into 4 main genotypes (Fig. 2a and Supplementary Figure 1). Most (83%) of the sequences belonged to the BA (BA9, BAC, BA10, BA, BA7, BA4, and BA6) genotype, followed by the CB1 genotype, SAB (SAB3 and SAB4) genotype and GB3 genotype with prevalence of 11%, 3.0%, and 0.7%, respectively. Genotypes BA9, BAC and BA10 were the main clusters of BA genotype with 68%, 6.2% and 6.2% in overall prevalence. The BA9 genotype, the largest cluster in this study, included sequences from 13 provinces from 2006 to 2014. The BA10 genotype clustered mostly with sequences collected from samples from Gansu in 2008 and in Beijing from 2009 to 2010. Three different lineages were detected for the BA9 as well as the BA10 genotypes (Fig. 2a and Supplementary Figure 1).

Forty-two Chinese BAC sequences were collected from 5 regions of China (all 6 regions except the Dongbei region) between 2008 to 2010, which constituted an independent lineage with 78% bootstrap value (Fig. 2a and Supplementary Figure 1). A total of 76 sequences from 6 regions clustered with the sequence of genotype CB1 (7 of these 76 sequences were designated as genotype GB5 in a previous publication<sup>17</sup>) and this genotype circulated in China between 2008 and 2013. HRSVB SAB3 strains were only found in Hong Kong SAR in 2000 and 2004, while SAB4 strains circulated in all regions except the Dongbei region during 2000–2001 and 2004–2010 seasons after this genotype was first detected in Jilin. Only 2 sequences from Hong Kong SAR and Chongqing in 2001 and 2010, respectively, were identified as GB3 genotype.

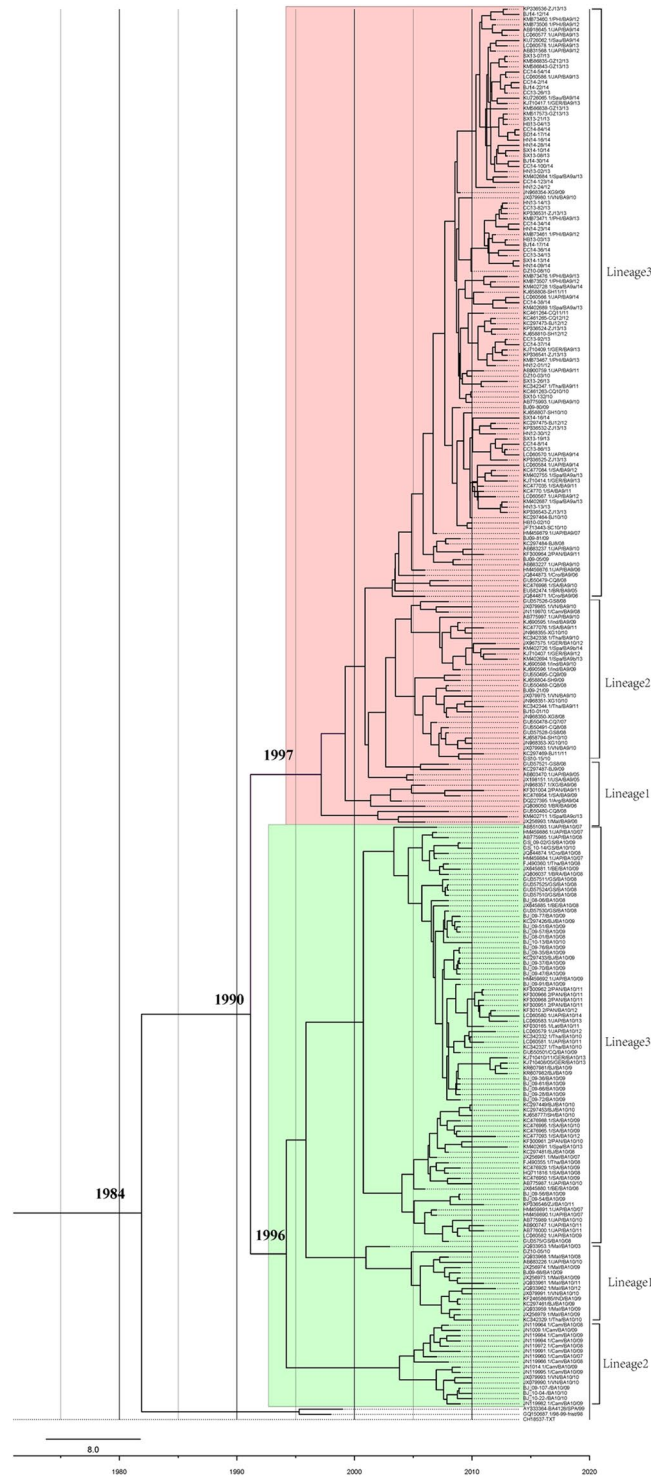
**Yearly and geographic distribution of HRSVB genotypes.** Genotypes SAB and CB1 were the predominant HRSVB genotypes in China before 2007 (Fig. 3). Following that year, the BA genotype became the main HRSVB genotype and circulated with the CB1 genotype simultaneously in China until 2013. The first genotype BA sample identified in this study was collected from Beijing in 2004. From 2008 to 2010, the BA9 genotype was the most prevalent genotype in China, co-circulating with a small number of HRSVB samples of the BA10, BAC, and CB1 genotypes. From 2012 to 2014, the number of the HRSVB sequences that clustered with the BA9 genotype increased substantially, with the prevalence of BA9 reaching almost 100% of HRSVB. In contrast, the BA10 genotype was found circulating intermittently in China in 2008–2011. It was most frequently detected in 2009 and 2010, but the number of BA10 samples gradually decreased in the following years to the level that no BA10 samples were detected in 2014 (Fig. 3).

**Phylogenetic analysis of genotype BA9 and BA10 sequences from China and other parts of the world.** The spread of the BA9 and BA10 genotypes has been a worldwide phenomenon in recent years. To study the relationship of the BA9 and BA10 genotypes from China with those from other parts of the world, 92 representative Chinese BA9 sequences were selected based on their yearly and geographic origins, as well as their nucleotide diversities to be analyzed in phylogenetic studies with 64 global BA9 reference sequences downloaded from GenBank. Sequences from the BA9 genotype could be divided into 3 potential lineages, named lineages









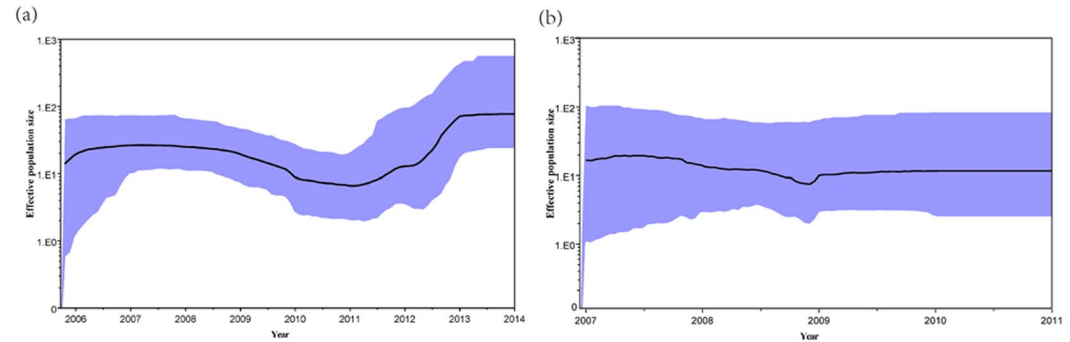
**Figure 4.** Maximum clade credibility tree of genotypes BA9 and BA10 of HRSVB. Sequences of the BA9 genotype were shaded in pink and those of the BA10 genotype were shaded in green.

between 2010 and 2014. It will be interesting to monitor if the BA10 genotype would reappear in the future or the BA9 genotype would circulate in China for a long period of time.

Phylogenetic analysis revealed that sequences of the BA9 genotype collected from China could be grouped into 3 lineages with sequences collected from other parts of the world. All of these 3 lineages could be detected in countries in Asia (e.g. China, Philippines, Vietnam, Japan, Saudi Arabia, Thailand and Malaysia), Europe (Spain, Brazil, Belgium, Germany and Croatia), South Africa and Panama from 2010 to 2014. For example, the sequences in lineage 3 from Japan<sup>29</sup>, Spain<sup>30</sup> and Germany<sup>31</sup> were named as lineage 1, BA9-A and Cl3, respectively; sequences in lineage 2 were named as BA9-B and Cl2 in Spain and Germany, respectively<sup>29–31</sup>. These data

Genotype	Number of sequences	Evolutionary rate (95% HPD) (substitution/site/year $\times 10^{-3}$ )	TMRCa(95%HPD) (year)
BA9	156(92CHN)	4.53 (3.4, 5.86)	1997 (1985, 2016)
	92CHN	4.38 (2.81, 6.12)	1996 (1986, 2017)
BA10	114(43CHN)	4.21 (2.94, 5.55)	1996 (1984, 2015)
	43CHN	4.13 (2.01, 6.63)	1997 (1992, 2016)
BA9/BA10	267(130CHN)	4.03 (2.99, 5.09)	1990 (1975, 2015)

**Table 2.** Evolutionary rates and TMRCa of genotypes BA9 and BA10. CHN: Chinese sequences; HPD: Highest probability density; TMRCa: Time to the most recent common ancestor.



**Figure 5.** (a) Bayesian skyline plot with 92 Chinese BA9 sequences from 2006 to 2014. (b) Bayesian skyline plot with 43 Chinese BA10 sequences from 2008 to 2011. The thick solid line is the estimated median, and the grey areas show the upper and lower bounds of the 95% HPD interval.

suggested that these 3 lineages emerged or circulated widely in Asian and European countries, raising the possibility that BA9 of these lineages were transmitted between China and other countries. Similarly, phylogenetic analysis also showed that the 43 sequences of Chinese BA10 genotype clustered into 3 lineages with sequences collected from other parts of the world. Of interest, lineage 3 of genotype BA10 appeared to have spread widely and Chinese sequences demonstrated high similarity with sequences from Japan<sup>16</sup>, Malaysia<sup>15</sup>, Thailand<sup>32</sup>, and India<sup>33</sup> suggesting the possibility of transmission of this lineage between China and neighboring countries.

Based on our analysis, the TMRCa of the BA9 or BA10 genotype was estimated to be 1996 or 1997, respectively. However, BA9 and BA10 genotypes were first detected in Japan in 2006 and 2007, indicating that both genotypes remained undetected for 10 years<sup>29</sup>. In addition, the BEAST analysis estimated that the evolutionary rates of the BA9 genotype ( $4.53 \times 10^{-3}$  substitutions/site/year) and the BA10 genotype ( $4.21 \times 10^{-3}$  substitutions/site/year) were not significantly different, and these evolutionary rates were similar to that reported by Japan analysis<sup>29</sup>. The BEAST analysis also suggested that the evolution rates of Chinese BA9 and BA10 genotypes were similar to those of the global BA9 and BA10 genotypes. Interestingly, the BSP analysis found that the effective population size of BA9 genotype increased from 2011 till 2014 suggesting that the genetic diversity of Chinese BA9 genotype had increased during this period of time. It remains to be determined if this increase in genetic diversity has contributed to the development of the different BA9 lineages in China. Although it is not clear if the evolution of the BA genotype enhances its spread, the BA genotype has spread widely all over the world during the past 15 years evolving into 12 genotypes<sup>19,27</sup>. It will be important to monitor if the BA9 genotype would continue its spread in the world, or if the BA9 lineage 2 or/and 3 would prevail in China for an extended period of time.

The BAc genotype was only identified in HRSVB from China; 42 BAc sequences clustered away from other genotype sequences with 78% bootstrap value. The BAc genotype was first identified in samples collected in the 2008/2009 and 2009/2010 epidemic seasons<sup>11</sup>. In this study, the BAc genotype was found to be circulating in China from 2008 to 2012, 2 years longer than previously reported. However, no BAc samples were detected in 2013 and 2014. Continued surveillance is necessary to monitor the reappearance of this genotype.

The CB1 genotype was first detected in Beijing in 2009, and the GB5 genotype was first detected in Chongqing in 2010<sup>11,17</sup>. In this study, we confirmed the finding of a previous report that the genotype GB5 samples collected in Chongqing<sup>17</sup> and the genotype CB1 samples collected in Shanghai<sup>23</sup>, Zhejiang<sup>24</sup> and Sichuan<sup>25</sup> were of the same genotype because they clustered together in our phylogenetic analysis. In addition, the sequences identified as CB1 in this study clustered into an independent branch away from any GB genotype reference sequences in the phylogenetic analysis.

There are some limitations of our study due to the difference of sample collection in each year or geographic location that may impact the estimation of the prevalence of HRSVB genotypes in China: the number of HRSVB samples collected before 2008 was small, the number of samples from each province was different each year, and samples were not collected from some provinces every year. In addition, this study was a retrospective study and the source of available samples was limited by the number of samples collected previously.

In conclusion, we found that there was a co-circulation of the BA with the non-BA genotypes in China between 2005 and 2013. The BA9 genotype emerged in 2006 and was the predominant HRSVB genotype circulating in China from 2008 to 2014. The genetic diversity of the BA9 genotype has increased from 2011 to 2014. Hence, a continuous and comprehensive surveillance is needed to monitor the circulation and evolution of the BA and non-BA genotypes to understand the HRSVB transmission in China in the future.

## Methods

**Ethics statement.** This study was approved by the second session of the Ethics Review Committee of the National Institute for Viral Disease Control and Prevention of the Center for Disease Control and Prevention (CDC) in China and the methods were performed in accordance with the approved guidelines. There was no human experimentation involved in this study; only nasopharyngeal precipitates were collected from patients with respiratory infections. Written Informed consents for the use of the clinical samples were signed by parents or guardians.

**Specimen collection.** Nasopharyngeal precipitates were taken from patients who were hospitalized with acute respiratory illness or outpatients with symptoms of respiratory infections in 1991, 2004 and from 2008 to 2014. This study was approved by 9 sentinel hospitals, from the Dongbei region (Jilin province), Huabei region (Beijing and Hebei provinces), Huadong region (Shandong province and Shanghai), Zhongnan region [Guangdong and Hunan provinces, and Hong Kong special administrative region (SAR)], and Xibei region (Gansu and Shaanxi provinces). Clinical samples were inoculated into HEP-2 cells for virus isolation before storing at  $-80^{\circ}\text{C}$ . Double-channel real-time reverse transcription-polymerase chain reaction (RT-PCR) was conducted to identify the HRSV-positive samples as well as to determine their HRSV subgroups<sup>34</sup>. To confirm the HRSV subgroups of the samples, the sequences were compared with HRSV sequences of known subgroups by the Basic Local Alignment Search Tool (BLAST) program available on the NCBI homepage (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

**RNA extraction, DNA amplification, and sequencing.** The QIAamp RNA mini kit (QIAGEN, Valencia, CA, USA) was used to extract the total RNA from clinical samples or the isolates according to the manufacturer's instructions. RT-PCR was performed using Takara one-step RT-PCR kit (TaKaRa Biotechnology, Dalian, China) to amplify the sequence of the region encoding the HVR2 of the G protein (637–968 nt) of HRSVB using forward primer GPB: 5'-AAGATGATTACCATTTTGAAGT-3', and reverse primer: F1: 5'-CAACTCCATTGTTATTTGCC-3'<sup>35</sup>. The PCR products were purified using a QIAquick Gel Extraction Kit (QIAGEN), and the purified PCR products were sequenced using an ABI Prism 3710xl DNA Analyzer at Majorbio Co., Ltd. (Beijing, China). The HRSVB sequences were edited using Sequencher software version 5.0 (Gene Codes, Ann Arbor, MI, USA).

The 365 sequences generated in this study are available in GenBank with accession numbers of DQ289648, DQ289649 and KX892710 to KX893072.

**Additional Chinese HRSVB sequences used in this study.** 332 Chinese HRSVB sequences that were available in GenBank on July 18, 2016 were downloaded to be analyzed together with the 365 HRSVB sequences generated in this study using the MEGA5.0 software.

**Phylogenetic analysis.** The HRSVB sequences were aligned with reference HRSVB sequences using the ClustalW program in the MEGA5.0 software ([www.megasoftware.net](http://www.megasoftware.net)). The phylogenetic trees were constructed using the neighbor-joining and the maximum composite likelihood methods within MEGA5.0 software, using 1000 replicates of bootstrap probabilities with a cut-off of  $\geq 70\%$  for evaluation of confidence estimates<sup>36,37</sup>.

**Evolutionary rate analysis and estimation of TMRCA.** The evolutionary rate and TMRCA were estimated by BEAST (Version 1.8.2)<sup>38</sup>. The jModelTest software (Version 0.1)<sup>39</sup> was used to calculate the best nucleotide substitution model and the GTR (general time reversible model) + G (gamma) were determined as the best substitution model for BA9 and BA10 genotype sequences. The uncorrelated exponential relaxed clock model was considered to be the best after the Bayes factor test of the strict, uncorrelated log-normal relaxed clock and the uncorrelated exponential relaxed clock. The Markov Chain Monte Carlo chains were run for 100,000,000, and 400,000,000 steps. After 10% burnin, a Maximum Clade Credibility tree was built by TreeAnnotator (Version 1.8.2) and edited by FigTree (Version 1.4.2). The BSP analysis was conducted using the BEAST 1.8.2 package.

## References

1. Pretorius, M. A. *et al.* The role of influenza, RSV and other common respiratory viruses in severe acute respiratory infections and influenza-like illness in a population with a high HIV sero-prevalence, South Africa 2012–2015. *J Clin Virol* **75**, 21–6 (2016).
2. Branche, A. R. & Falsey, A. R. Respiratory syncytial virus infection in older adults: an under-recognized problem. *Drugs Aging* **32**, 261–9 (2015).
3. Naorat, S. *et al.* Hospitalizations for acute lower respiratory tract infection due to respiratory syncytial virus in Thailand, 2008–2011. *J Infect Dis* **208**(Suppl 3), S238–45 (2013).
4. Feng, L. *et al.* Viral etiologies of hospitalized acute lower respiratory infection patients in China, 2009–2013. *PLoS One* **9**, e99419 (2014).
5. Rudan, I. *et al.* Causes of deaths in children younger than 5 years in China in 2008. *Lancet* **375**, 1083–9 (2010).
6. Zhang, Y. *et al.* Burden of respiratory syncytial virus infections in China: Systematic review and meta-analysis. *J Glob Health* **5**, 020417 (2015).
7. Mufson, M. A., Orvell, C., Rafnar, B. & Norrby, E. Two distinct subtypes of human respiratory syncytial virus. *J Gen Virol* **66**(Pt 10), 2111–24 (1985).
8. Peret, T. C. *et al.* Circulation patterns of group A and B human respiratory syncytial virus genotypes in 5 communities in North America. *J Infect Dis* **181**, 1891–6 (2000).



9. Venter, M., Madhi, S. A., Tiemessen, C. T. & Schoub, B. D. Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes. *J Gen Virol* **82**, 2117–24 (2001).
10. Baek, Y. H. *et al.* Prevalence and genetic characterization of respiratory syncytial virus (RSV) in hospitalized children in Korea. *Arch Virol* **157**, 1039–50 (2012).
11. Cui, G. *et al.* Genetic variation in attachment glycoprotein genes of human respiratory syncytial virus subgroups a and B in children in recent five consecutive years. *PLoS One* **8**, e75020 (2013).
12. Eshaghi, A. *et al.* Genetic variability of human respiratory syncytial virus A strains circulating in Ontario: a novel genotype with a 72 nucleotide G gene duplication. *PLoS One* **7**, e32807 (2012).
13. Arnott, A. *et al.* A study of the genetic variability of human respiratory syncytial virus (HRSV) in Cambodia reveals the existence of a new HRSV group B genotype. *J Clin Microbiol* **49**, 3504–13 (2011).
14. Blanc, A., Delfraro, A., Frabasile, S. & Arbiza, J. Genotypes of respiratory syncytial virus group B identified in Uruguay. *Arch Virol* **150**, 603–9 (2005).
15. Etemadi, M. R., Sekawi, Z., Othman, N., Lye, M. S. & Moghaddam, F. Y. Circulation of human respiratory syncytial virus strains among hospitalized children with acute lower respiratory infection in malaysia. *Evol Bioinform Online* **9**, 151–61 (2013).
16. Dapat, I. C. *et al.* New genotypes within respiratory syncytial virus group B genotype BA in Niigata, Japan. *J Clin Microbiol* **48**, 3423–7 (2010).
17. Ren, L., Xiao, Q., Zhou, L., Xia, Q. & Liu, E. Molecular characterization of human respiratory syncytial virus subtype B: a novel genotype of subtype B circulating in China. *J Med Virol* **87**, 1–9 (2015).
18. Jinhua, S. *et al.* Emergence of ON1 genotype of human respiratory syncytial virus subgroup A in China between 2011 and 2015. *Sci Rep* **7**, 5501 (2017).
19. Trento, A. *et al.* Ten years of global evolution of the human respiratory syncytial virus BA genotype with a 60-nucleotide duplication in the G protein gene. *J Virol* **84**, 7500–12 (2010).
20. Trento, A. *et al.* Natural history of human respiratory syncytial virus inferred from phylogenetic analysis of the attachment (G) glycoprotein with a 60-nucleotide duplication. *J Virol* **80**, 975–84 (2006).
21. Zhang, R. F. *et al.* Human respiratory syncytial virus in children with acute respiratory tract infections in China. *J Clin Microbiol* **48**, 4193–9 (2010).
22. Zhang, Z. Y. *et al.* Genetic variability of respiratory syncytial viruses (RSV) prevalent in Southwestern China from 2006 to 2009: emergence of subgroup B and A RSV as dominant strains. *J Clin Microbiol* **48**, 1201–7 (2010).
23. Liu, J. *et al.* Genetic variation of human respiratory syncytial virus among children with fever and respiratory symptoms in Shanghai, China, from 2009 to 2012. *Infect Genet Evol* **27**, 131–6 (2014).
24. Yu, X. *et al.* Human respiratory syncytial virus in children with lower respiratory tract infections or influenza-like illness and its co-infection characteristics with viruses and atypical bacteria in Hangzhou, China. *J Clin Virol* **69**, 1–6 (2015).
25. Hu, P. *et al.* Alternate circulation and genetic variation of human respiratory syncytial virus genotypes in Chengdu, West China, 2009–2014. *J Med Virol* **89**, 32–40 (2016).
26. Trento, A. *et al.* Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. *J Gen Virol* **84**, 3115–20 (2003).
27. Agoti, C. N., Gitahi, C. W., Medley, G. F., Cane, P. A. & Nokes, D. J. Identification of group B respiratory syncytial viruses that lack the 60-nucleotide duplication after six consecutive epidemics of total BA dominance at coastal Kenya. *Influenza Other Respir Viruses* **7**, 1008–12 (2013).
28. Ohno, A. *et al.* Genetic characterization of human respiratory syncytial virus detected in hospitalized children in the Philippines from 2008 to 2012. *J Clin Virol* **57**, 59–65 (2013).
29. Nagasawa, K. *et al.* Molecular evolution of the hypervariable region of the attachment glycoprotein gene in human respiratory syncytial virus subgroup B genotypes BA9 and BA10. *Infect Genet Evol* **36**, 217–23 (2015).
30. Gimferrer, L. *et al.* Circulation of a novel human respiratory syncytial virus Group B genotype during the 2014–2015 season in Catalonia (Spain). *Clinical Microbiology and Infection* **22**, 97.e95–97.e98 (2016).
31. Tabatabai, J., Prifert, C., Pfeil, J., Grulich-Henn, J. & Schnitzler, P. Novel respiratory syncytial virus (RSV) genotype ON1 predominates in Germany during winter season 2012–13. *PLoS One* **9**, e109191 (2014).
32. Auksornkitti, V. *et al.* Molecular characterization of human respiratory syncytial virus, 2010–2011: identification of genotype ON1 and a new subgroup B genotype in Thailand. *Arch Virol* **159**, 499–507 (2014).
33. Sahu, M., Shukla, M. K. & Barde, P. V. Molecular characterization of human respiratory syncytial virus detected from central India. *Journal of Medical Virology* **89**, 1871–1874 (2017).
34. van Elden, L. J. R. *et al.* Applicability of a Real-Time Quantitative PCR Assay for Diagnosis of Respiratory Syncytial Virus Infection in Immunocompromised Adults. *Journal of Clinical Microbiology* **41**, 4378–4381 (2003).
35. Zhang, Y. *et al.* Genetic variability of group A and B human respiratory syncytial viruses isolated from 3 provinces in China. *Arch Virol* **152**, 1425–34 (2007).
36. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–25 (1987).
37. Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–9 (2011).
38. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214 (2007).
39. Posada, D. jModelTest: phylogenetic model averaging. *Mol Biol Evol* **25**, 1253–6 (2008).

## Acknowledgements

We thank the staffs in the 9 sentinel hospitals from Jilin, Beijing, Hebei, Shandong, Shanghai, Guangdong, Hunan, Gansu and Shaanxi provinces who performed specimen collection, viral cultures and clinical information data collection. This work was supported by the Key Technologies R&D Program of National Ministry of Science (2013ZX10004-202, 2012ZX10004201-003) and Non-clinical laboratory service agreement with AbbVie Pharmaceutical Trading (Shanghai) Co. Ltd.

## Author Contributions

W.X., Y.Z. conceived and designed the experiments. J.S., A.C., Y.H., L.S., Z.Y., X.X., J.S., H.W., C.M., P.Y., and Q.L. performed the experiments. J.S., Y.Z., R.Z., W.X. analyzed the data. J.S., Y.Z., T.N. and W.X. wrote the main manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-17055-0>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017